Best
Available
Copy

AD-A281 448

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

2. REPORT DATE

3. REPORT TYPE AND DATES COVERED
FINAL 01 Feb 93 TO 31 Jan 94

SITUATIONAL AWARENESS FOR COMPLEX SYSTEM OPERATIONS

5. FUNDING NUMBERS
F49620-93-1-0132
61102F
2313
BS

6. AUTHOR(S)
Dr Richard Gilson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Univ of Central Florida
CAHFA
Orlando, FL 32816-0390

8. PERFORMING ORGANIZATION REPORT NUMBER
AEOSR-TR- 94 0418

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NL
110 DUNCAN AVE SUITE B115
BOLLING AFB DC 20332-0001
Dr John F. Tangney

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release;
distribution unlimited.

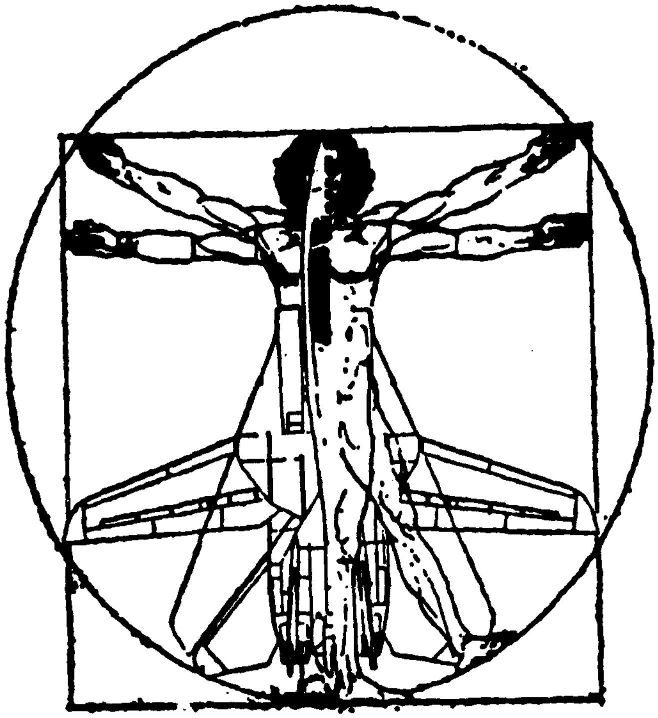
12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Decisions, whese appropriateness depends on knowledge and rules, are exected more or less effectively depending on the psychomotor skills of the operator. If these are sequential rather than parallel processes, and it seems to me that they must be even though they may be carried out very quickly, then in some cases we may be able to infer decision from action, and gestalt from descision, as long as we understand that the cognitive centroid of the individual operator is idiosyncratic and unknowable to some degree. Training can help to improve perception: it can also help to standardize the decisions taken in a given situation. The comprehension and integration of sensed data can also be improved by training, practice and criticism. The changes brought about by carefully targeted training can be observed and can also help us to underlying processes.

14. SUBJECT TERMS	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
(U)	(U)	(U)	(U)



Situational Awareness in Complex Systems

Edited by:

Richard D. Gilson Daniel J. Garland Jefferson M. Koonce

Aviation Human Factors Series

94 7 8 052

Situational Awareness in Complex Systems

Proceedings of a CAHFA Conference

Aviation Human Factors Series from
The Center for Applied Human Factors in Aviation (CAHFA):

- **Verification and Validation of Complex Systems: Additional Human Factors Issues (1993)**
- **Human Factors and Advanced Aviation Technologies (1993)**
- **Situational Awareness in Complex Systems (1994)**
- **Human Factors Certification of Advanced Technologies (1994)**

Situational Awareness in Complex Systems

Proceedings of a CAHFA Conference

Edited by

Richard D. Gilson, Ph.D.
Department of Psychology
University of Central Florida
Orlando, Florida

Daniel J. Garland, Ph.D.
Center for Aviation/Aerospace Research
Embry-Riddle Aeronautical University
Daytona Beach, Florida

Jefferson M. Koonce, Ph.D.
Center for Applied Human Factors in Aviation
University of Central Florida
Orlando, Florida

Embry-Riddle Aeronautical University Press
Daytona Beach, Florida USA 1994

Proceedings of the Center for Applied Human Factors in Aviation (CAHFA) Conference on *Situational Awareness in Complex Systems*, Orlando, FL, February 1–3, 1993.

Notice

This document is disseminated under the sponsorship of the U.S. National Center for Atmospheric Research, the U.S. Air Force Office of Scientific Research, the U.S. Federal Aviation Administration, the U.S. Army Research Institute for the Behavioral and Social Sciences, Embry-Riddle Aeronautical University, and the University of Central Florida in the interest of information exchange. Opinions expressed in this document do not necessarily reflect official policy or approval of the sponsors or the United States Government. The United States Government assumes no liability for the contents or use thereof.

The United States Government does not endorse products or manufacturers. Trade or manufacturer's names appear herein solely because they are considered essential to the object of this document.

Technical Editing by

L. W. Hennessy
Center for Aviation/Aerospace Research
Embry-Riddle Aeronautical University
Daytona Beach, Florida

Graphic Designs by

Nancy Rahn-Virden
Embry-Riddle Aeronautical University
Daytona Beach, Florida

ISBN 1–884099–02–5

For information contact:
Embry-Riddle Aeronautical University Press
600 South Clyde Morris Blvd.
Daytona Beach, FL 32114-3900

Table of Contents

Air Traffic Control

Weather Services

Aerospace Operations

Emergency Management Systems

Team Awareness

Addendum

Closing Comments

Foreword

On behalf of the Center for Applied Human Factors in Aviation (CAHFA) and the consortium institutions, University of Central Florida and Embry-Riddle Aeronautical University, it is our distinct pleasure to offer these proceedings as an enduring part of CAHFA's international symposium on situational awareness. The symposium was held in Orlando, Florida, on February 2-4, 1993 and attracted an assembly of the world's experts in the area of situational awareness.

Since situational awareness often seems to mean something different to almost everyone, the objective of the conference was to begin to scope the concept, delineating what is known and what needs to be known by identifying gaps in knowledge and future needs. And, as most practitioners are interested in what one can *do* with situational awareness, the conference's main thrust sought real world recommendations for enhancing time compressed assessments of complex systems operating in dynamic environments. Participants examined theoretical and applied issues of situational awareness including operational definitions, research methodologies, measurement techniques, training and selection, team behavior, and designs that enhance human capabilities.

The conference produced some notable agreements: first, certainly that the concept of situational awareness is important and worth pursing. It was observed that the loss of situational awareness has resulted in some of the most publicized human factors accidents, e.g., Three Mile Island, the USS Vincennes, and Eastern Airlines flight 401 that descended into the Florida Everglades. Further, we were reassured as to the value in seeking an understanding of situational awareness by successes like the emergency landing of United Airlines flight 232 at Sioux City, Iowa with the many lives saved and the recent thwarting of several intricate international terrorist plots in the United States.

An accord emerged among authors and participants alike that the concept of situational awareness is a difficult one, indeed. It was affirmed that situational awareness is dynamic, that it can be fragile--taking time to rebuild, and that it must include some kind of an organized and internally consistent mental model. Further, the construct of situational awareness goes well beyond any immediate system awareness to include events in the environment surrounding the system as well as what is brought to bear by one's past experiences.

How situational awareness might be used was of great interest; for instance, how to anticipate future events based on currently developing situations, thereby preventing potential mistakes or accidents in the making, or by seizing opportunities that might otherwise be lost. Discussions indicated that one viable application may be to provide intervention early in the chain of events for most effective prediction and control of developing situations. Furthermore, heightened situational awareness may help to cope with the unexpected by enhancing ongoing assessment and developing flexible variations to unanticipated events. Perhaps the major frustrations that emerged were the lack of data in support of theory, the lack of well defined methodologies for research, and the lack of agreement as to measures. On the other hand, it is encouraging to accept that these frustrations provide the challenges for future research and opportunities for future symposiums.

Rather than posit our own definition of situational awareness, we will allow our contributing authors that challenge. It is clear that situational awareness goes well beyond performance alone. The following papers suggest that situational awareness is much more cognitive in nature than performance; a cognizance of one's surroundings, an alertness in drawing inferences from what one experiences.

Richard D. Gilson
Daniel J. Garland
Jefferson M. Koonce

Acknowledgments

The development and publication of a volume of this nature is not without a tremendous amount of dedication and hard work from a group of quality individuals. The editors would like to acknowledge the work of those individuals and organizations who made the conference and publication of this volume possible.

We must thank our sponsors, without whom the conference could not have taken place. The sponsors for the conference included:

- U.S. National Center for Atmospheric Research
- U.S. Air Force Office of Scientific Research
- U.S. Federal Aviation Administration
- U.S. Army Research Institute for the Behavioral and Social Sciences
- Embry-Riddle Aeronautical University
- University of Central Florida

We extend a fervent thank you to all the conference participants and authors for their enthusiastic contributions and for their boldness in addressing such a difficult issue. We must thank an excellent conference staff who worked hard before, during, and after the meeting. The conference was managed with the outstanding and tireless work of Mr. Eric V. Gruber, CAHFA's administrator. Eric was bolstered by an exceptional staff, particularly Mr. Joseph Pittle with UCF's Center for Continuing Education, who was on site at the conference virtually day and night, and Mr. Srinath Sogal at the Center for Aviation/Aerospace Research at Embry-Riddle Aeronautical University who contributed a yeoman's effort in compiling these proceedings. We owe a significant debt to L. W. Hennessy for his exceptional technical editing of these proceedings, and to Nancy Rahn-Virden for her outstanding graphics which significantly contributed to the conference program and proceedings. Finally, a heartfelt thank you goes to the graduate and undergraduate students who put in untold hours to meet the omnipresent on-the-spot needs of such a conference.

We are forever grateful for the support of these many individuals and sponsoring organizations in the publication of this volume.

Richard D. Gilson
Daniel J. Garland
Jefferson M. Koonce

Introduction

- Yours, Mine, and Ours: Some Observations on the Metaphysics of Situational Awareness
- An Introduction to the Concept of Situational Awareness

Yours, Mine, and Ours: Some Observations on the Metaphysics of Situational Awareness

Lloyd Hitchcock

Hitchcock & Associates, Inc.

I would first like to express my thanks to the organizers of this Conference for providing me with the opportunity to exchange views with those of you who have attended these proceedings. As someone who is deeply concerned with the process of Air Traffic Control, I am constantly involved with, or perhaps more appropriately, embattled against situational awareness issues and all of their many operational and cognitive ramifications. However, in no way could I, or would I, lay claim to being a preeminent researcher in this area or even an unusually knowledgeable student of the contemporary literature in the field. At the same time, as an avid, perhaps even desperate at times, user of the results of situational awareness investigations, it is my hope that I may provide you with some insights, and their corollary challenges, which might prove to be useful to us all, researcher and user alike.

Situational awareness, or "SA" as I am sure I shall refer to it on occasion during the rest of our discussion, would seem to be in the larger sense a significant component of practically everything that is done by any and all sentient creatures. SA is frequently associated with such activities as Air Combat and Air Traffic Control; perhaps because of its obvious criticality to the successful performance of those functions. This is not to diminish the importance of adequate situational awareness to such other activities, as broken-field running, highway or urban driving, baby-sitting (my wife and I have raised four children separated by 6 years), etc. Indeed, it is hard to envision a circumstance in which an individual might be without situational awareness. While such tragedies as the oft cited "controlled flight into terrain", midair collisions, and fuel starvation downings of "lost" pilots serve as ample evidence of the frequency with which we fallible humans are plagued by inadequate, incorrect, and/or inappropriate SA templates, I would question the contention that those involved in such unfortunate circumstances as these were ever totally without at least some concept of the world about them. I am, of course, well aware of the anecdotal evidence cited regarding ATC controllers who, as a consequence of extreme workload, distraction, or both, have "lost the bubble" and appear to be without even a working hypothesis of the nature of the world around them. However, I would argue strongly that far, far more problems of this kind are caused by faulty situational awareness templates (cognitive maps if you would prefer) than by no situational awareness at all.

I have been discussing situational awareness and have yet to define what it is about which we are attempting to exchange ideas. As Sarter and Woods (1991) have stated, while "situational awareness has become a very popular phrase in the aviation domain, its use is

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

most often based on an intuitive, not necessarily appropriate understanding". These authors state further that "situation awareness ... has become a ubiquitous phrase and the frequent topic of research projects even without consensus on its meaning or much knowledge about existing problems which need to be addressed". The fact that situation awareness remains largely undefined is certainly not the result of a lack of trying; perhaps it is more likely the result of an over enthusiastic plethora of attempts. Whitaker and Klein (1988) have defined situational awareness as a pilot's "knowledge about his surroundings in light of his mission goals". Tenney, et al. define situational awareness as "the up-to-the-minute cognizance required to operate or maintain a system". Hartman, et al., define SA as a "skill" and Forrester (1978) defines it as a "sixth sense". By postulating that situational awareness is "trainable", Endsley; Garland, et al.; Kass, et al.; and Swartz have all, more or less, endorsed the concept of situational awareness as a skill. Sarter and Woods question the exact nature of the distinction that is made, at times, between situational awareness and the companion concept of "mental models". Both Kieras and Bovier and Garland and his co-authors admonish us to be sensitive to the need to separate the "process" of situational awareness from the "products" that are the results of that process. Neely also has approached situational awareness as a process and has given us two models of the information acquisition and processing involved in the development of situational awareness; one that is primarily automatic and subconscious and a second mode which is "controlled, contemplative, and acutely conscious".

As concluded by Harwood, as well as many others, the primary elements or components of situational awareness are usually couched in terms of the quantifiable spatial aspects of what these authors label as the "pilot's knowledge of his location in space and of the spatial relationship between objects" within that space (Harwood, et al., 1988). As detailed in his comprehensive reviews of the measurement of situational awareness (1988, 1991), Major Fracker (USAF) has identified the absence of disparity between objective measurements of the spatial parameters in a given situation and a test subject's mental model reconstruction of that configuration as the primary index of the relative "goodness" of that subject's situational awareness. As have others, Endsley (1988, 1989) has endorsed the use of "snapshots" of such measures of disparity as an index of the "pilot's mental model of the world around him" and demonstrated their utility by successfully correlating the goodness of the "situational awareness", implied by the magnitude of such measures, to success in aerial combat. I would, in no way, attempt to argue against the importance of such a conceptual definition of "situational awareness". Not only have these researchers demonstrated both the utility and validity of such measures but, for one such as myself who professes an interest in the realm of the air traffic controller, to discount the importance of any type of four dimensional spatial construct would border on madness. However, the direct comparison of a subject's mental model with that of the experimenter's has, at least to me, some disturbing aspects as a research protocol. For one thing, this protocol implies a confidence and certainty in the independent reality and the uniform, existential, validity of our "objective and quantitative" measures of those elements of our world which we select to use to define the situation about which we wish to become aware. This is, in and of itself, a questionable assumption. Not only is there a significant body of philosophical discourse which declares such a contention to be logically impossible but we have, readily at hand, more objective evidence which places that view in serious doubt. For example, within the world of the Air Traffic Controller, a dominant variable is "altitude". An aircraft's altitude is certainly quantifiable in terms of specific measures within an agreed upon frame of reference (that is, if we chose to ignore the problems known to be associated with the distinctions between "altitude above ground (AGL)" and "above sea level (MSL)" and make sure we don't worry about the differences

associated with the use of inches of mercury, millibars, and hectopascals as the metric of reference). However, if we do indeed use an agreed upon metric, logic would tell us that it should be as easy to convey situational awareness about the altitude, either present or desired, of one aircraft, as it would be to communicate the altitude of any other aircraft.

However, we find that appears not to be the case. Don George, of NASA's Aviation Safety Reporting System (ASRS) program, analyzed hundreds of reports of altitude "busts" by commercial airline crews. His findings revealed a significant amount of confusion between 10,000 and 11,000 feet (Figure 1) and the frequency of busts associated with all other altitude pairings which he analyzed. George cites the fact that it appears to be "easy to confuse "one-one-thousand" and "one-zero-thousand". However, I personally fail to see that compelling a difference between that particular pairing and the calls, for example, for "two-one-thousand" and "two-zero-thousand". While George's findings may simply reflect nothing more than the fact that far more controller/pilot exchanges are tied to the 10 and 11 thousand foot pairing than to any of the other alternative combinations, they should give us pause before we accept, totally without question, the assumption of intra-observer equivalence between the situational awareness elements, even those that are quantitative, which comprise the respective mental models of the situation in which our subject's are operating. Certainly we are fortunate that there is a significant correlation between our uniquely modified "life spaces". If this were not the case, any such activities as Air Traffic Control or Coordinated Aerial Combat would be totally impossible. However, I would contend that, in our understandable fixation upon the objectively quantifiable as the independent and dependent variables in our research, we may have run the risk of ignoring the unique subjective aspects of our respective "worlds" and thus may have failed to account for a significant portion of the variance associated with our results. If, as George's work would appear to confirm, there are significant intra-observer disparities within the attributes of specific situational elements, it should come as no surprise that there is a potential for inter-observer non-agreements as well. Indeed there is a well founded philosophical basis to support the assumption that such differences will, in fact, exist. The Dutch Topologist, Luitzen Brouwer, went so far as to deny even the existential reality of numbers themselves declaring that numbers exist only as the products of the thought processes of the mathematicians who are thinking about them. Thus, there are as many sets of numbers as there are mathematicians.

In his somewhat theological, and almost fully philosophical novel "Illusions", Richard Bach, of "Jonathan Livingston Seagull" fame, portrays himself flying wing through the Midwest with another pilot; both selling airplane rides for $3.00 per person. Bach's new friend, Donald Shimoda, has admitted to being a "reluctant messiah" and is attempting to tutor Bach into achieving a similar status. One afternoon while they are lying in a hayfield under the wings of their aircraft, Bach remarks to Donald:

> "Donald, I have come to the conclusion that you just don't live in this world".
> Shimoda replies, "Of course not, Can you tell me one person that does?" Bach's immediate answer is, "ME, I live in this world".
> "Excellent," Shimoda replies. "Remind me to buy you lunch today... I marvel at the way you never stop learning." Bach continues with, "Of course I live in this world. Me and about four billion other people..."

At that point, Donald informs Bach that he has "just blown lunch" and goes on to chide him for thinking, even for a moment, that he lives in the same world as a stockbroker whose life space has just been devastated by the recent SEC ruling requiring mandatory review of any investor portfolio showing a shareholder loss of more than fifty percent. Donald goes on to point out that Bach's "world" is far, far removed from that of a professional tournament chess player whose attention has been riveted on what would have been, for that time, the equivalent of Fischer, Spasky, and Serajevo. Nor are his concerns those of, say, an aspiring watercolorist, a commercial fisherman, etc.

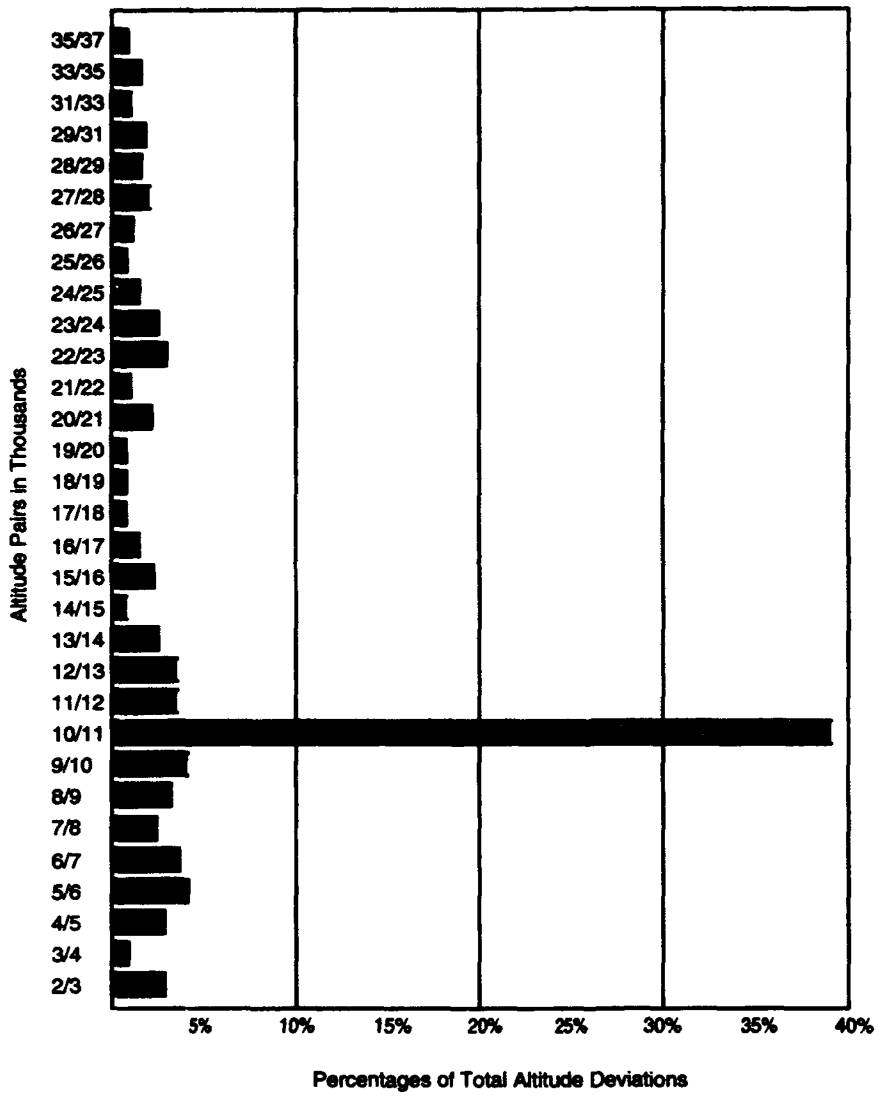


Figure 1. Percentages of Altitude Deviations by Altitude Pairing

Bach's unique world, as defined by Shimoda, is a 1929 Fleet Biplane landed in a "hayfield in Maitland, Ohio" with his "major life's priorities farmer's permission, people who want ten minute airplane rides, [and] Kinner aircraft engine [parts and] maintenance". Donald goes on: "Are you standing there ... telling me that four billion people do not live in four billion separate worlds?". We all have personal examples of how circumstances have changed our "worlds". In my own case, I clearly remember how my purchase of a Cessna 150, a light aircraft covered with thin aluminum skin, forever changed, dramatically, my visceral response to the meteorological term "hail".

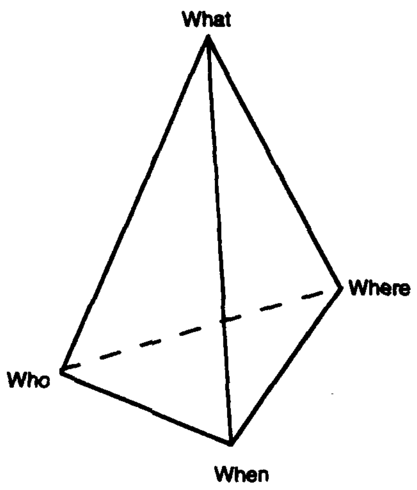


Figure 2. The "pseudo" pyramid of Harwood, Barnett and Wickens.

Since I am well aware of the validity of Kaiser's Second Law which states "Never open a can of worms unless you plan on going fishing", I will not try to proceed further out onto the potentially hazardous metaphysical "ice" of relative existentialism. My objective in this allegorical discussion was, simply, to put before us, for further consideration, the rather obvious, but sometimes overlooked, fact that the mental models we label as "situational awareness" are not only determined by the quantifiable attributes of the physical world but are also filtered through the subjective elements of each individual's own particular "world". Perhaps the most significant implication of this philosophical recognition of the separate uniqueness of our individual "worlds" is the warning it carries with it regarding what we face when we try to introduce one observer's mental model of a given situation into that of another.

Another, perhaps better, way to describe what I am attempting to convey would be to utilize the geometric representation of situational awareness developed by Harwood, Barnett, and Wickens. These authors distilled four, primary parameters to describe a model of situational awareness; "where", "what", "when", and "who". They went on to graphically organize these elements into what they called a "pseudo" pyramid (Figure 2). While this diagram serves as a compelling and convenient graphic summary of the four-dimensional SA space, it carries with it implications which can potentially affect our understanding of

situational awareness which I find disturbing. This construct postulates direct linkages between the four components; the classic approach of those oriented toward the objectively quantifiable. However, I would contend that this model is not an empty polygon as shown, but is, rather, filled with the uniqueness of the individual observer. Those of you who are not aware that I am a product of a good Skinnerian background might not recognize what a wrenching declaration that is for me to make. However, it seems more fitting to me that information about the "what" of a situation is not passed directly to either the "who", "where", and/or "when" components but is, rather, detoured through a nexus, a "Cognitive Centroid" (make a note, you encountered that particular term here first) which sits in the middle of the polygon (Figure 3) and represents the sum total of the memories, linguistic peculiarities, background information data bases, etc., which characterize the uniquely personal aspects of each observer's "world".

I would further contend that when one person attempts to communicate with, or more importantly, tries to create in the mind of another a reproduction of the situational awareness they have of their own "world", this information transference does not take place between the points which define Harwood's polygon but (Figure 4) is, rather, mediated between the centroids of the respective worlds of each individual involved. This double filtration process cannot help but introduce a significant source of variance between the measurements of the parameters taken in one situation and the assessment of the same data sampled from another's construct of the same situation no matter how precise our quantification methodologies might be.

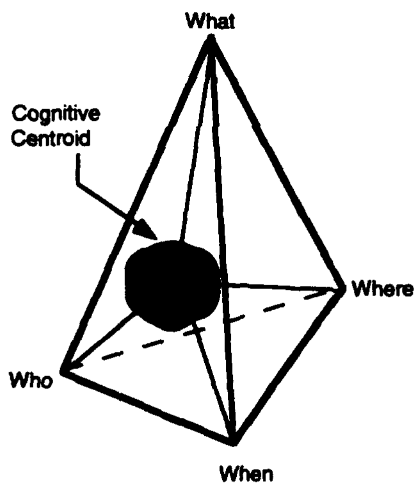


Figure 3. The "Cognitive Centroid."

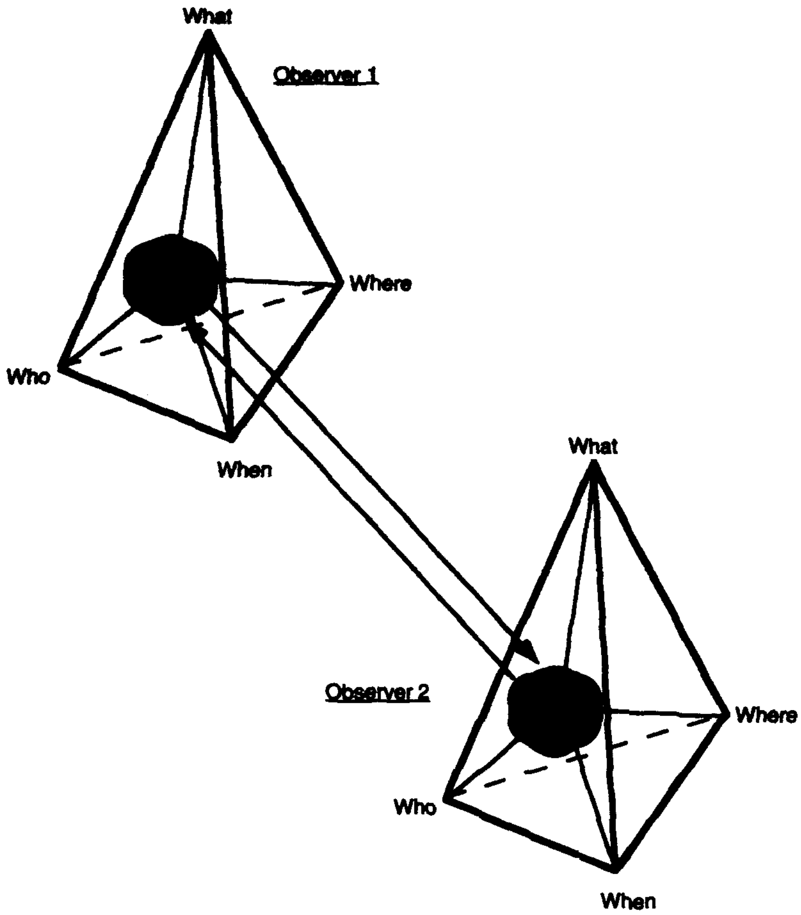


Figure 4. *The transfer of information.*

As an illustration of how two people can have virtually identical mental models of the locations and spatial relationships defining a particular three dimensional environment and yet achieve far different success in communicating their respective situational awareness to another, consider the following scenario (Figure 5). This map shows a small, hypothetical, segment of rural terrain in the Ozark mountains of southwestern Missouri where I grew up. You are at the "star" where you have turned in to a filling station to ask directions to the home of friends, Dr. and Mrs. Wilson, who have moved to a farm in the area. You ask the attendant how to reach the Wilson's farm and are pleased to find that he, indeed, knows them and knows where they live. He gives you the following set of "Ozark type" directions:

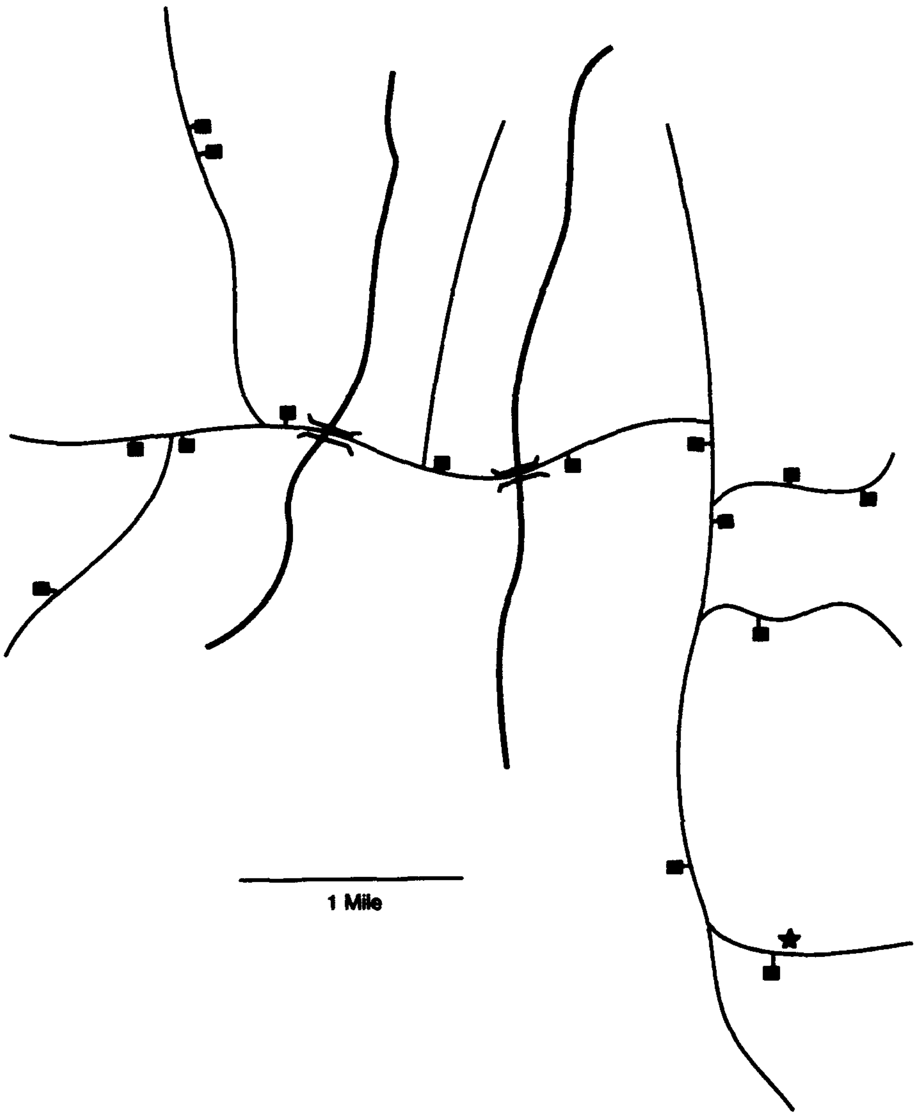


Figure 5. Map of hypothetical setting.

Direction Set 1

1. Take the Marshfield Road.

2. Turn at the old Hasting's place. You can't miss it, they're "goat-in" it off.

3. Stay on Pine Road until just past James Creek and turn on Old Mill Road just on the other side of the Owens' driveway.

4. Doc Wilson bought the old Peterson place about two years ago from old man Jimmy Lee's daughter; the one who got divorced and moved to Tulsa.

I am sure that none of you, provided with these directions could manage to find your friends' new home. By the way, for the very few of you who might not be familiar with the term "goat-in it off", I will explain. Often, when a hill farmer decides to clear land, particularly acreage that has been tilled before, it is a common practice to loose a herd of goats onto the land. Given no other food, the goats will quickly browse off the weeds and undergrowth down to bare ground making the soil far easier to plow for replanting. The importance of this fact, as a component of the direction giver's instructions, is that, to one familiar with this practice, the sight of a goat grazed hillside is absolutely unmistakable. Fortunately, another customer at the filling station has overheard the exchange between you and the attendant and comes to the rescue. Using more conventional, and far more generic, terminology, he provides you with an alternative set of directions to locate the farm you are seeking:

Direction Set 2

1. When you leave the station, go left and, in about two hundred yards, take a right at the cross roads.

2. Continue to go North on this road for almost three miles and take the first road to the left.

3. In just over one mile, turn right just after you cross the second creek.

4. You will reach the Wilson's in about a mile. They live in the first house on your right.

Looking at this second set of directions side by side with the map, I would expect that virtually all of you have, by now, have successfully located the Wilson's farm. Does this mean then the second set of directions reflects a superior mental model of this particular spatial situation? It certainly proved to be a better template for communicating "awareness" from the direction given to you. When we look at the area map annotated by these directions (Figure 6), the route to our destination would appear to be both clear and unambiguous.

However, let's go back to the first set of directions and see how they work when they are superimposed upon the basic outline map (Figure 7). If we look at this set of directions, armed with the knowledge base of the gentleman who gave them to us, we have no more difficulty reaching the Wilson's than we did with the more conventional guidance. Indeed, in some respects, this set is perhaps even more efficient than the other. It transfers the needed information with the same number of specific items: four. It uses unique knowledge to eliminate ambiguity rather than spatially referenced directions. For example, if you know where the Hasting's place is, you do not need to be told which way to turn; there is only one choice. In the same vein, if you know which creek is "James Creek", you do not need to count waterways.

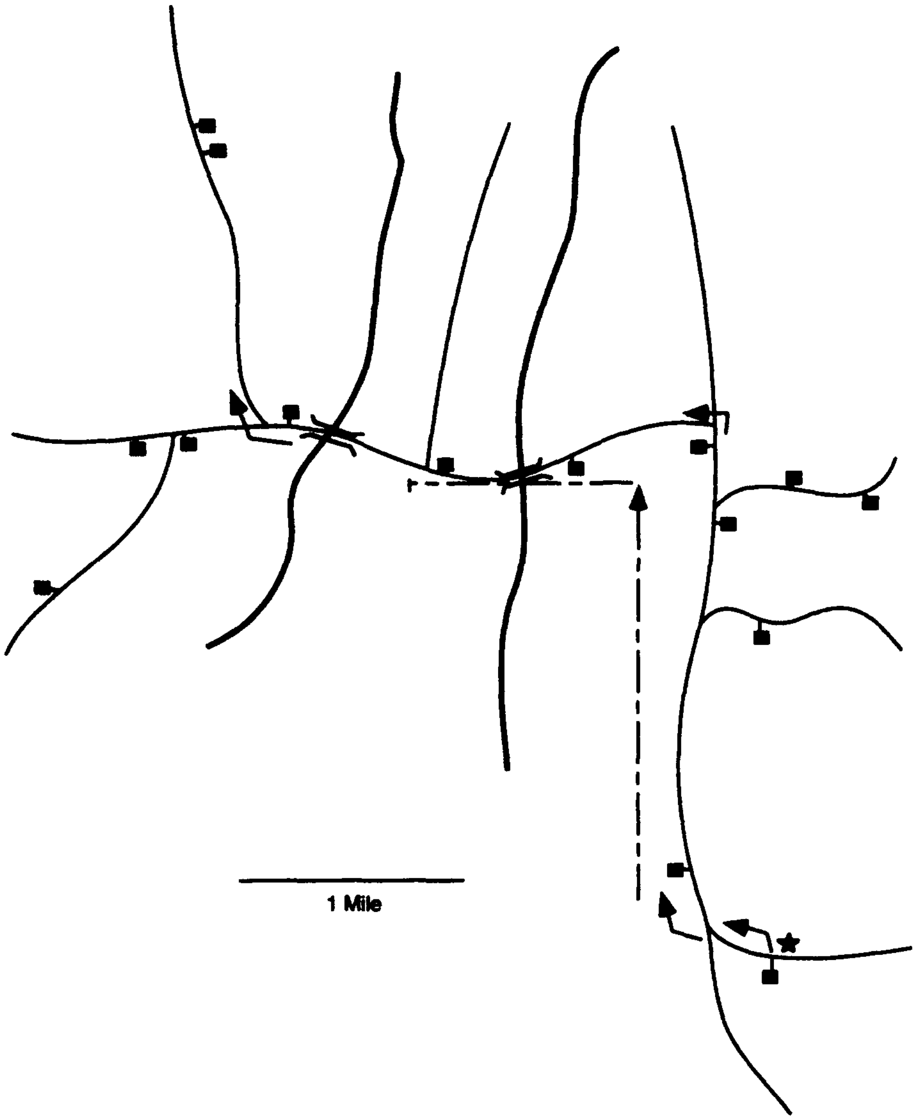


Figure 6. Annotated map.

Besides, look how information rich this colloquial set of directions is. You not only know how to get to the Wilson's farm but you know that the farm previously owned by the Hasting's is undergoing preparations to expand its tillable acreage, you also know who sold the Wilson's the farm they are now living in, and you are now up-to-date on the life and times of the Peterson's daughter. Not necessarily a bad communications "grade" for the process involved in one person's creation of what we might call a "mental model" for the use of another.

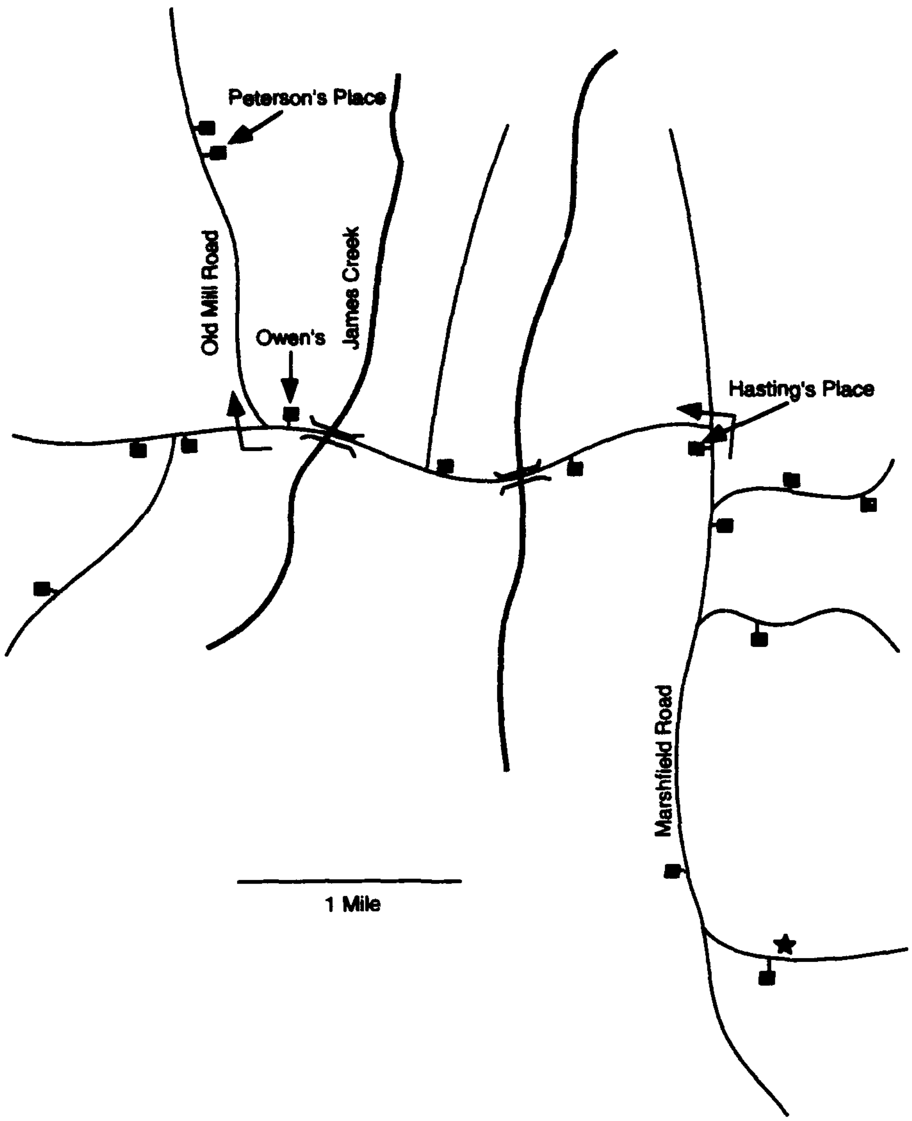


Figure 7. Initial directions superimposed on the basic outline map.

In case you might have the impression that such convoluted attempts at the transference of situational data, while interesting and, possibly, even humorous, could never be a part of the technologically sophisticated world in which we routinely interact with each other, may I add one more anecdote of a personal nature. The first time I flew into Philadelphia International, I was directed to a straight in approach to Runway 17 and asked to "report passing GE". I had to come back with the fact that I was not familiar with the area and did not know the General

Electric Plant. The controller's reply was "Roger – Zero Five Fox, report crossing 69th Street". The best I could do at that point was give a position report at "three miles out" and that had to do. This is not a unique circumstance. A VFR approach to Runway 36 at Washington National is often accompanied with a request to report "passing the powerhouse". This is not as challenging as the identification of the "GE" plant or "69th Street" posed by the Philadelphia controller.

The "Powerhouse", located at the northern edge of Old Town Alexandria, is rather readily distinguishable by its multiple smokestacks and the block long mountain of coal beside it. There is another example of two perceived situations in conflict reported by Bill Richard's in the Aviation Safety Reporting System Bulletin "Directline". Just as a three-engine wide-body was starting to level off at its assigned altitude of 41 thousand feet, it experienced a compressor stall and had to shut down one of its engines. The flight crew "...advised center [that they were] descending, had shut down an engine, and need[ed] 24,000 feet". The controller cleared the aircraft for a descent to flight level 370. Twice the flight crew advised that they had to "get down to a lower level". However, the controller had traffic at 35,000 feet and was unable to approve their request without a high risk of loss of separation. The flight crew kept repeating its request for a lower altitude and the controller kept repeating that he was "unable". Other than the fact that the flight crew lacked any knowledge of the conflicting traffic beneath them, had we taken a "snapshot" reading of this situation as understood by both the flight crew and the controller, there would have been little or no disparity between the four-dimensional representations of the geographic and/or spatial parameters which they would have produced. This particular ATC problem arose because there were different expectancies superimposed upon the controller's and the flight crew's spatial templates. The controller had heard, and believed, the oft repeated claims that this type of three-engine heavy aircraft could "fly all day" on two engines and was treating the flight crew's need to descend as a routine change in altitude request. Of course, the flight crew knew that their aircraft could not maintain altitude on two engines at their current flight level; the laws of physics mandated that they come down to at least 24,000 feet whether the controller approved their descent or not. However, the crew was equally aware that, were they able to reach their one-engine-out operational altitude, they could continue their scheduled flight without incident. Therefore, they did not feel compelled to declare an in-flight emergency. I am sure that an analysis of their cockpit voice recorder would have revealed a number of choice comments about controllers who expected them to ask their passengers to pull up as hard as they could on their seat belts in order to maintain assigned altitude. In this instance, a summary comparison of the objective, quantifiable aspects of these participants' situational awareness templates would, in all likelihood, have failed to reveal a problem of any kind between these two "worlds".

The importance of the "observer" as an intervening variable in the understanding of situational awareness is not a fresh concept for which I would, even for a moment, attempt to take credit. Bolstad (1991) has included "personality factors" as one of the six attributes which his search of the literature showed to be correlated with situational awareness. Endsley's inclusion of Chase and Simon's concept of "expertise" as an important component of the definition of situational awareness certainly presupposes a significant individual contribution to the SA process. We can address this problem of observer specific SA templates by revising our attribute definitions to be more inclusive. However, if we follow that path, we run the risk, cited by Sarter and Woods, of becoming "too general to really help in understanding situational awareness". Or, as Rouse and Morris (1986) pointed out "to find out that humans have to know 'something' in order to perform their tasks... [would be] no

great stride for science". Woods (1986) has counseled us that "little progress has been made with respect to important issues such as supporting the acquisition of information in... data rich environments."

Thus, as I see it, we face three separate but interrelated challenges in our discussion of SA. The first is to better understand how each of us creates his or her own unique mental model of our surroundings: our situational awareness. Secondly, we should consider how each of us uses our individual gifts of awareness most effectively. And, lastly, we most efficiently and accurately communicate our own unique situational awareness between one another.

I literally stand in awe at the intellectual audacity we display when we propose to address the challenges we have placed before us for this conference. We are not only taking on some of the most demanding psychological and psychometric issues of subject matter definition, experimental design, and measurement methodology imaginable, but are doing so in the face of overwhelming historical precedent that says that, philosophically at least, what we are proposing to do can not be done at all. It is my sincere hope that, throughout our discourse we will be able to reach some level of peaceful accommodation in our collective "War of Our Worlds" and that the contents of "My World" will have joined the attributes of your own unique universe, and that somewhere "out there", suspended in the cognitive void which separates us, we will have succeeded in creating a space, an "Our World." Where we can reach a new level of understanding: a heightened awareness of this situation in which our common interest has brought us together.

References

Bach, R. (1977). *Illusions; the adventures of a reluctant Messiah*, Delacorte Press Inc., (ISBN: 0-440-04318-2).

Bolstad, C. A., (1990). *Evaluation of pilot attribute measures.* Northrop Aircraft Division Report. Hawthorne, CA.

Endsley, M. R., (1988). *Situation awareness global assessment technique (SAGAT).* Paper presented at the National Aerospace and Electronic Conference (NAECON), Dayton, OH.

Endsley, M. R., (1988) Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors 32nd Annual Meeting.* Santa Monica, CA.

Endsley, M. R., (1989) *A methodology for the objective measurement of situation awareness.* Paper presented at the AGARD Symposium on Situation Awareness in Aerospace Operations, Copenhagen.

Forrester, L., (1978) *Fly for your life*, 2nd ed., Banram Books, Inc., New York, NY.

Fracker, M. L., (1988) A theory of situation awareness: Implications for measuring situation awareness. In *Proceedings of the Human Factors Society 32nd Annual Meeting*, Santa Monica, CA., pp. 102-106.

Fracker, M. L., (1989) Attention allocation in situation awareness. In *Proceedings of the Human Factors Society 33rd Annual Meeting, Denver, Co.*, pp. 1396-1399.

Fracker, Major M. L., (1991) Measures of situation awareness. Technical Report, Armstrong Laboratory, Crew Systems Directorate, Wright-Patterson AFB, OH.

Garland,D. J., Stein, E. S., Blanchard, J. W., & Wise, J.A. (1992) Situational awareness in the future air traffic control environment, *Proceedings of the 37th Air Traffic Control Association Conference*, Atlantic City, NJ.

George, D. (1991). *One zero ways to bust an altitude*. In ASRS Directline, NASA, Moffett Field, CA., Fall.

Hartman, B. O., & Secrest, G. E. (1991) Situational awareness is more than exceptional vision. In *Aviation, Space, and Environmental Medicine*, November, pp. 1084-87.

Harwood, K., Barnett, B., & Wickens, C. (1988). Situational awareness: A conceptual and methodological framework. In F. E. McIntire (Ed.), *Proceedings of the 11th Symposium of Psychology in the Department of Defense* (pp. 316-320). Colorado Springs, CO, US Air Force Academy.

Kass, S. J., Herschler, D. A., & Companion, M. A. (1990) Are they shooting at me?: An approach to training situational awareness. *Proceedings of the Human Factors Society 34th Annual Meeting*, Santa Monica, CA, 1352-1356.

Kieras, D. E., & Bovier, S. (1984) The role of a mental model in learning to operate a device. *Cognitive Science*, 8, pp. 255-273.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: roles of inhabitionless spreading activation and limited capacity attention. *J. Exp. Psychol*, 106, 226-254.

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: prospects and limits in the search for mental models. *Psychological Bulletin*, 100, 349-363.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology*, 1(1), 45-47.

Schwartz, D. (1989) Training for situational awareness. In *Proceedings of the Ohio State 5th International Symposium on Aviation Psychology*, April, Columbus, OH.

Tenney, Y. J., Adams, M. J., Pew, R. W. Huggins, A. W. F., & Rogers, W. H. (1992) A principled Approach to the measurement of situation awareness in commercial aviation. NASA Contractor Report No. 4451, Langley Research Center.

Whittaker, L. A., & Klein, G. A. (1988). Situation awareness in the virtual world: Situation assessment report. *Proceedings of the 11th Symposium of Psychology in the Department of Defense*, Colorado Springs, CO., US Air Force Academy.

Woods, D. D., (1986) Paradigms for intelligent decision support. In E. Hollnagel, G. Mancici, & D. D. Woods (Eds.) *Intelligent decision support in process environments*. North-Holland, New York, NY., pp. 153-173.

Woods, D. D. (1988). Coping with complexity: The psychology of human behavior in complex systems. In L. P. Goodstein, H. B. Anderson, & S. E. Olsen (Eds.) *Tasks, errors, and mental models*. Taylor & Frances, London, pp. 128-148.

An Introduction to the Concept of Situation Awareness

Richard A. Pew

Bolt, Beranek and Newman Inc.

Introduction

My purposes in this paper are to provide a definition of Situational Awareness (SA) that can serve across the disciplines that wish to use the term, to describe a framework based on psychological theory and data that addresses situational assessment, to discuss the process of achieving SA and to raise some important defining questions that I hope will be further addressed by other papers in this volume.

When pushed for a one sentence definition of SA, I call on Endsley (1988): "The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and projection of their status in the near future." This definition captures the notion of spatial awareness that is important to piloting examples, but also leaves room for awareness of systems' states for non-moving systems or the status of systems that do not relate to the elements in the airspace, such as a nuclear power plant safety parameter display system. It acknowledges the fact that it is not just data from the environment that matter, but also the interpretation of those data utilizing the crewmembers' knowledge and experience. Finally, it includes the notion that effective SA is useful for anticipating what is likely to happen in the future as well as knowledge of the immediate present.

There is particular interest in SA these days because advanced technology, the introduction of extensive automation into today's systems is alleged to be taking it away. Those of us working on human interaction with systems are advocating that "Human-Centered Design" will bring it back and it is clear that successful performance depends on having it. We therefore need to find out (1) how to select individuals who have it, (2) how to train individuals to improve their SA skills, and (3) how to design systems to make it easier for crewmembers to achieve SA. Besides, SA has become the Buzz-Word of the 90's!

Formal Definition of Situational Awareness

In order to adequately define SA we need to understand what we mean by a "situation" and we need to know what it is about situations of which we must be aware. It is also of

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

interest to catalog where that information and knowledge come from. I adopt Table 1 as a working definition of a situation:

Table 1. Definition of a Situation

A situation is a set of environmental conditions and system states with which the participant is interacting that can be characterized uniquely by a set of information, knowledge and response options.

However, the concept of a situation is meaningful, according to this view, only if we can define a discrete and innumerable set of them; that is, that the awareness requirements can be broken discretely into packets, each associated with a set of system states. This ability to partition situations implies that, while the environment is more or less continuously changing with time, only some of the changes are large or severe enough to create a changed situation from the perspective of the crewmember. Examples of such changes that are severe enough to redefine the SA might be: a forest fire that has run out of its firebreak, a power plant that transitions form start-up to full power or an aircraft autopilot disengagement, either expectedly or unexpectedly.

The second part of the definition requires that a "situation" have associated with it the information and knowledge that we are calling awareness. Table 2 shows the elements that need to be included and Table 3 lists the informational sources that the crewmember has to draw on to achieve SA.

Having defined situations and the to be defined components of awareness is still not enough because we must also define the requirements for SA that are presented by a given situation. Otherwise we have no standard against which to judge how successful a crewmember has been in achieving SA. Our thinking about this has led to the consideration of an ideal awareness, the obtainable ideal as well as the actual SA achieved.

Table 2. Elements of Awareness, Given the Situation

- Current state of the system (including all the relevant variables).
- Predicted state in the "near" future.
- Information and knowledge required in support of the crew's current activities.
- Activity Phase
- Prioritized list of current goal(s)
- Currently active goal, subgoal, task
- Time
- Information and knowledge needed to support anticipated "near" future contexts.

The *ideal* is that SA that is defined by experts evaluating the requirements at leisure or even after the fact. It includes both the information and knowledge requirements. The *obtainable ideal* is that subset that is actually available for the crewmember to acquire. When defining the obtainable ideal we assume the availability of well designed information resources and take into account the fact that what is practically available is constrained by

expected human cognitive abilities. It does not seem fair to evaluate actual SA achievement by a crew member with respect to a goal that is not practically achievable. Both the ideal and the achievable ideal are assessed independently of crewmember performance. However, the *actual* SA must be inferred from measurement or observation. It is the difference between actual SA and ideal SA that creates the space for evaluating individual differences in the ability to achieve SA and for developing training and design alternatives contributing to improved SA.

Table 3. Information Resources Contributing to Awareness

- Sensory information from the environment
- Visual and auditory displays
- Decision aids and decision support systems
- Extra- and intra-crew communication
- Crewmember background knowledge and experience

We are currently examining ways of evaluating ideal and achievable ideal SA on the basis of an analytic simulation model of the environment and the crewmembers' performance in that environment.

Situational Assessment

It is useful to distinguish the process of achieving SA, which we will refer to as Situational Assessment, from the product that results from that process. It takes active effort to achieve SA. The process is best regarded as a special case of mental workload that competes with other aspects of task performance. The product is the resulting awareness that we can measure. It is the product of SA that governs task performance. If we are trying to develop procedures for assessing individual differences in SA or if we are developing training procedures, we are mainly interested in SA as process or situational assessment. However, when we are interested in system design to promote SA, we need to obtain an understanding of both the product and the process. In this section, we suggest a psychological framework for characterizing situational assessment. The material that follows in this section has been condensed from Tenney, Adams, Pew, Huggins, and Rogers, (1992).

In a broad perspective, although the aircrew spend much of their time in routine, repetitive activities, at any moment in time a number of different, potentially knowledge intensive and procedurally complex tasks may demand attention. Each of these tasks is usually triggered by a stimulus event, such as an ATC communication or a warning light, and, in order to achieve adequate situational awareness, may require that the crew initiate additional information seeking behavior. The aircrew needs to be prepared for these by having an adequate mental picture of the situation and knowing exactly what additional information is needed. For each such alerting signal or task, the crew must determine its relevance, its procedural implications and its urgency. For the experienced crew, often these events call forth highly practiced patterns and result in "automatic" responses. If not, assessing their significance may require access to the full range of human memory structures to retrieve associated data and knowledge necessary for deciding on a course of action.

The Available Information

In carrying out these tasks, the crewmember must be attentive to numerous sources of information. There is an ongoing stream of sensory information from the environment as well as visual and auditory displays, manuals, checklists, and communication between crewmembers or from external sources

In some circumstances, the crewmember can choose to ignore or attend to incoming information on the basis of simple dimensions, such as location or modality. Location provides information about the function of certain inputs. Location also provides information about the criticality of certain inputs (e.g., aircraft warnings are listed before cautions which are listed before advisories). Criticality is also specified by modality (e.g., the distinctive sound that signals a warning).

While the relevance of certain inputs can be determined easily, the relevance of others is not signaled by either source or modality, but depends on its relationship to some particular aspect of the situation. For example, in the case of a commercial aircraft on an approach to landing, the significance of a radio message cannot be determined without further processing: A message from ATC to reduce speed and watch for traffic would certainly be relevant; A pilot report from the preceding pilot would be relevant if it contained information about a potential hazard; a communication addressed to another pilot would be irrelevant, unless the pilot were nearby and it described conditions at the relevant airport; and a radio message that began "Standby for SIGMET" (a special weather advisory) would be relevant if it pertained to the right part of the country.

To summarize, one factor that determines whether and how easily an input can be processed has to do with the structure of the information. The significance of certain inputs is directly specified by superficial, attention getting features while that of others requires deeper processing and further information seeking.

Human Information Processing Constraints

Recent work on memory and attention suggests that another factor affecting the processing of an input is the extent of the competing demands on the pilot's attention at any given moment. The bottom line is that human information processing capabilities are not well suited to a multiplicity of tasks. Thoughtful attention is modular. People can consciously think about only one thing at a time; as a result, they do not handle interruptions and distractions very well. Research has shown that even when an operator is faced with as few as two tasks and the "tasks" consist of nothing more than the detection or recognition of simple signals, the requirement to divide or switch attention between them may result in a significant loss in sensitivity or time that can be allocated to either (Broadbent, 1957; Schneider & Detweiler, 1988; Swets, 1984).

In addition, research indicates that mental shifts between topics or semantic domains require measurable time and effort and are prone to certain classes of biases and errors (Anderson & Pitchard, 1978; Bower, 1982; Sanford & Garrod, 1981; Schank, 1982). To the extent that incoming information is unrelated to the task in which the pilot is concurrently engaged, its interpretation must involve considerable mental workload. The more time and effort the crewmember invests in its interpretation, the greater its potential for blocking notice or proper interpretation of other available data. The less time and effort the crewmember invests in its interpretation, the greater the likelihood of misconstruing its implications.

A theory about memory and attention by Sanford & Garrod (1981) helps to clarify the cognitive demands on the crewmember, even though, interestingly, it was formulated to explain text comprehension.

Sanford and Garrod (1981) have theorized that an individual's active memory consists of two bins, *explicit* and *implicit* focus. *Explicit focus* corresponds roughly to what is conventionally labeled as the "short-term store." It is the working space within which a crewmember integrates new information with information that is already known. At any given moment, explicit focus contains a tightly limited number of interrelated tokens of (or pointers to) larger knowledge structures in long-term memory. Although the contents of explicit focus are regulated more or less like a push down stack, their maintenance of any given token depends not only on the recency with which it has been activated by the input from the environment, but also on its implicit relevance to the current interpretive stream.

Implicit focus, in contrast, subtends the full blown representation of the situation that is partially represented in explicit focus. Information relevant to the knowledge in implicit focus can be brought to the interpreter's attention with neither the speed nor the obliqueness of reference that suffices for information in explicit focus. On the other hand, such information can be interpreted far more quickly and with far less support than information that is unrelated to the contents of explicit focus.

To support these active memories, Sanford and Garrod suggest that the crewmember's latent (currently inactive) memory is also sectioned into two bins. The first, *long-term episodic memory,* contains a complete record of the knowledge structures that have been built or accessed in the course of the current system scenario. Meanwhile, *long-term semantic memory* contains a person's lifetime accumulation of knowledge in general. Knowledge in either of these latent memories can be brought to consciousness only given considerable effort or strong cueing.

Thus, for example, a pilot will readily notice and respond to changes in glideslope indication that he is tracking in the course of landing. Events that pertain to the task but not to the particular aspect of the task with which a person is engaged are also expected to be interpreted relatively quickly and cogently as they will map onto knowledge in implicit focus. Thus, for example, even while tracking the glideslope, the pilot may be readily alerted to changes in engine noise that are consistent or inconsistent with landing experience. In contrast, when the interpretation of an event requires consideration of knowledge in latent memory, the probability or effort associated with its proper processing will depend on such factors as the transparency of its significance and the time available for working on it: When very close to touchdown, for example, the pilot will be relatively unprepared to receive and interpret unrelated communications.

The Northwest Airlines accident at Detroit Metropolitan Airport in 1988 illustrates the difficulty a pilot faces in discriminating relevant from irrelevant information when the workload is high. The aircraft took off without setting its flaps and crashed. Although the crew had begun the pre-flight checklist properly, they were interrupted by ATC before verifying the status of the flaps. They might still have resumed the checklist routine prior to take off, but other issues usurped their attention: they were confused as to which taxiway to use, the runway direction had just been changed, and the weather and runway conditions were not provided until the aircraft was already taxiing. Although, with proper handing, the aircraft could have become airborne with the flaps mispositioned, the crew had been given a windshear alert. When the problem with the flaps expressed itself during take off, the symptoms were interpreted – and responded to – as thought they were caused by windshear.

As Figure 1 illustrates, this misinterpretation is understandable in light of the salience of windshear in the pilot's memory and the similarity of the symptoms in the two cases.

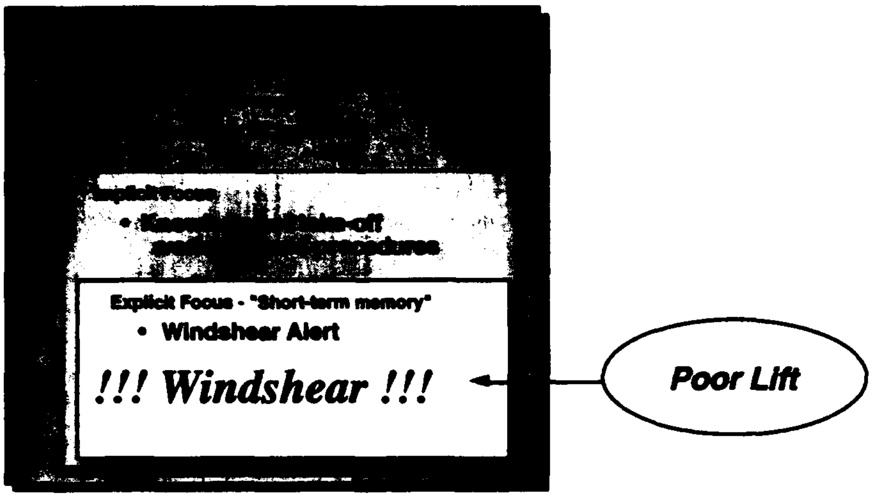


Figure 1. Sanford and Garrod Memory Representation Associated With the Northwest Airlines Crash.

A Critical Question

Having gained a perspective on the way in which SA is achieved, I now turn to a critical question that the SA researcher must answer before we can decide that it is truly worthwhile to commit significant resources to research on SA.

This question concerns the generality of the concepts associated with SA. Either (1) there are design principles and general individual abilities to be discovered that are applicable across specific devices and tasks, or (2) every task and device is unique and it only makes sense to select, train and design to promote the crew's performance with specific display configurations and response requirements.

If I may exaggerate only slightly, when one attempts to review the SA literature, one finds two types of studies: those that seek general measures of SA, frequently subjective measures, and apply them generically across a range of tasks, and those that address a specific design or training question for which it has been alleged that SA is an important criterion for success. These latter studies develop specific performance measures to address the specific design or training questions. For them, the concept of SA is not particularly important. Flack(1993), supporting that position, has argued eloquently that SA is a little like "the emperor's new clothes," in that there probably is not such a thing as general SA ability and that what we are really seeking is simply good performance.

For the moment, I come down in between. I believe it is not possible to define SA independently of an understanding of the SA requirements of context-specific situations. But I also believe that there may be skills in doing this that are more general and that are worthy of evaluation. Nevertheless, the emphasis in our work has been on developing objective measures that attempt to partition out that aspect of overall performance that might be called SA and that are specific to the context in which we wish to assess the concept.

Conclusions

There are situational awareness design principles, training opportunities and general individual abilities to be discovered that are applicable across specific devices and tasks. While they are not independent of the situations they support, we can think of SA as a general process supporting specific situations.

However, before leaping at these research opportunities, there is an urgent need to lay the defining ground work as a community so that common objectives can be sought and achieved. The operational definition of situational awareness, the product, its assessment, and the process, challenge our understanding of the fundamental mechanisms of cognitive performance.

References

Anderson, R.C., & Pitchard, J.W. (1978) Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior. 17,* 1-12.

Bower, G.H. (1982) Plans and goals; in understanding episodes. In A. Flammer & W. Kintsch (Eds.) *Discourse Processing*. New York: North-Holland.

Broadbent, D.E. (1957). "A mechanical model for human attention and immediate memory." *Psychological Review, 64,* 205-215.

Endsley, M.R. (1988). "Design and evaluation for situational awareness enhancement." *Proceedings of the Human Factors Society 32nd Annual Meeting.*

Flack, J.M. (1993) "Situation awareness: The emperor's new clothes." (Unpublished manuscript)

Sanford, A.J., & Garrod, S.C. (1981) *Understanding Written Language.* New York: Wiley & Sons.

Schank, R.C., Collins, G.C., Davis, E., Johnson, P.N., Lytinen, S., & Reiser, B.J. (1982). What's the point? *Cognitive Science, 6,* 255-275.

Schneider, W., & Detweiler, M. (1988). The role of practice in dual-task performance: Toward workload modeling in a connectionist/control architecture. *Human Factors, 30,* 539-566.

Swets, J.S. (1984). Mathematical models of attention. In R. Parasuraman & R. Davies (Eds.). *Varieties of attention.* New York: Academic Press.

Tenney, Y.J., Adams, M.J., Pew, R.W., Huggins, A.W.F., & Rogers, W.H. (1992). *A principled approach to the Measurement of situation awareness in commercial aviation.* NASA Contractor Report 4451.

Theory and Measurement

- Situation Awareness in Dynamic Human Decision Making: Theory
- Situation Awareness is Adaptive, Externally-Driven Consciousness
- Situation in Mind: Theory, Application and Measurement of Situational Awareness
- Situation Awareness in Dynamic Human Decision Making: Measurement

Situation Awareness in Dynamic Human Decision Making: Theory

Mica R. Endsley

Texas Tech University

This paper presents a theoretical model of situation awareness including its role in the dynamic decision making process. The model shows the relationship between operator goals, situation awareness and mental models in the selection of operator actions, and their role in directing the acquisition and interpretation of information from the environment to form situation awareness. Included are a general definition of the construct and a discussion of the impact of attention, workload and stress on operator situation awareness, along with a discussion of the role of expertise and automaticity. A model of decision making in dynamic, complex environments is presented, describing the role of situation awareness in such contexts. The types of errors which can occur in situation awareness are presented along with implications of this taxonomy for error reduction efforts.

Introduction

The focus of problems confronting human factors practitioners has continued to grow during the past 50 years. In addition to optimizing human performance in tasks which are primarily physical or perceptual, increasingly human behavior involving highly complex cognitive decision making tasks must be dealt with. As technology has evolved, many complex, dynamic systems have been created which tax the abilities of the human to act as an effective, timely decision maker in operating these systems. The operator's *situation awareness*, loosely defined as an internal model of the environment, will be presented as a crucial construct upon which decision making and performance in such systems hinge.

In this paper, I will strive to show (a) the important role that situation awareness plays in decision making in dynamic environments and the utility of using a model of decision making which takes situation awareness into account, and (b) a theory situation awareness, which has remained a somewhat enigmatic concept. True situation awareness, it will be shown, involves far more than merely being aware of numerous pieces of data. It also requires a much further level of situation understanding and comprehension and a projection of future system states in light of the operator's pertinent goals. As such, situation awareness presents a

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

level of focus which goes beyond traditional information processing approaches in attempting to explain human behavior in operating complex systems.

Situation awareness (SA) can be shown to be important in a variety of contexts which confront human factors practitioners.

Aircraft

In perhaps the area with the longest history, situation awareness was recognized as a crucial commodity for crews of military aircraft as far back as World War I (Press, 1986). SA has grown in importance as a major design goal for civil, commercial and all types of military aircraft, receiving particular emphasis in recent years (Federal Aviation Administration, 1990; Intraflight Command, Control and Communications Symposium, 1986). In the flight environment, the safe operation of the aircraft in a manner consistent with the pilot's goals is highly dependent on a current assessment of the changing situation, including details of the aircraft's operational parameters, external conditions, navigational information, other aircraft, and hostile factors. Without such assessment (which needs to be both accurate and complete), the aircrew will be unable to effectively perform their functions. Indeed, as will be discussed further, even small lapses in SA can have catastrophic repercussions.

Air Traffic Control

In a related environment, air traffic controllers are called upon to sort-out and project the paths of ever increasing numbers of aircraft in order to insure goals of minimum separation and safe, efficient landing and take-off operations. This taxing job relies upon the situation awareness of the controllers who must maintain up-to-date assessments of the rapidly changing location of each aircraft (in three-dimensional space) and their projected locations relative to each other, along with other pertinent aircraft parameters (destination, fuel, communications, etc.).

Large Systems Operations.

The operators of large, complex systems such as flexible manufacturing systems (FMS), refineries and nuclear power plants must also rely on up-to-date knowledge of situational parameters to effectively manage. In their tasks, the state of numerous system parameters must be observed, the patterns among them revealing clues as to the functioning of the system and predictions of process state changes (Wirstad, 1988). Without this understanding and prediction, human control could not occur in an effective manner.

Tactical and strategic systems

Similarly, fire fighters, certain police units and military command personnel rely on situation assessment to make their decisions. They must ascertain highly critical features of widely varying situations to determine the best course of action. Inaccurate or incomplete

SA in these environments can lead to devastating loss of life, such as in the case of the U.S.S. Vincennes, where incorrect SA concerning an incoming aircraft (due to confusing identification signals and a lack of direct information on changes in altitude), led to the downing of a commercial airliner and subsequent loss of all aboard. From reports of the accident (Klein, 1989a), it appears that it was the decision makers' SA that was in error (hostility of the incoming aircraft), not the decision as to what to do (if hostile, warn-off and then shoot-down if not heeded). This is an important distinction that highlights the criticality of SA in dynamic decision making.

Other

Many other every-day activities call for a dynamic update of the situation to function effectively. Walking or driving in heavy traffic and operating heavy machinery surely call for SA. Roschelle and Greeno (1987) report that experts in solving physics problems rely on the development of a situational classification. Many other tasks that people are involved in everyday rely on SA. As humans typically operate in a closed-loop manner, it can be shown that input from the environment is almost always necessary. As the tasks being performed become more and more complex, involving more numerous, detailed and inter-related data, a more detailed mental representation of that environment becomes necessary for performance.

(It is also worth noting at this point that the partial or full automation of systems in these environments will not remove the need for situation awareness as, in most cases, human operators will still be needed to oversee and interact with the systems. Quite the opposite, the trend towards automation may actually serve to increase the need for SA, as operators who are no longer actively working with information may be less likely to maintain SA, resulting in problems both in effectively monitoring the process and in taking over operations when necessary (Endsley, 1987b).)

In common to all of these areas is the fact that operators must strive to make complex decisions on the basis of the state of an ever-changing dynamic system and environment. Theories of decision making in the literature are inadequate for explaining operator behavior in these environments as, almost without exception, they have been developed on the basis of decision making for static tasks (e.g., choosing an apartment, career, automobile, etc.). In the real world, however, people often function as decision makers in *dynamic* tasks. Their tasks differ significantly from static tasks in that (a) many decisions are required across a fairly narrow space of time, and (b) the tasks are dependent on an ongoing, up-to-date analysis of the environment as a primary input to the decision making process. A key feature of skilled decision making in a dynamic environment is that the decision process is almost entirely driven by a conceptual assessment of the existing situation, hence the importance of situation awareness.

Dreyfus (1981) presented a treatise which emphasized the role of situational understanding in real-world, expert decision making, building upon the extensive works of deGroot (1965) in chess, Mintzburg (1973) in managerial decision making, and Kuhn (1970) in the field of science. In each of these three diverse areas, the experts studied use pattern-matching mechanisms to draw upon long-term memory structures that allow them to quickly understand a given situation. They then adopt a course of action which corresponds to that situational understanding. Hinsley, Hayes and Simon (1977) have found that this situation classification can occur almost immediately, or as Klein (1989b) has pointed out, can involve some effort to achieve. Researchers in many other areas have also found that expert decision

makers will act first to classify and understand a situation before proceeding with decision selection (Klein, 1989b; Klein, Calderwood, and Clinton-Cirocco, 1986; Lipshitz, 1987; Noble, Boehm-Davis, and Grosz, 1987; Sweller, 1988). Klein's (1989b) studies of fire ground commanders found that a conscious deliberation of solution alternatives was rarely observed in these cases. Rather, the majority of the time, the experts focused on classifying the situation, such classification immediately yielding the appropriate solution from memory. As decision makers will select actions on the basis of their perceptions of the situation, accurate and complete situation awareness is necessary for effective decision making in a dynamic environment. (While much of this work emphasizes the decision processes of experts, it can be seen that even novices in dynamic situations must focus a considerable amount of their effort on assessing the state of the environment as input to their decisions.)

Given that situation awareness plays an important role in dynamic decision making environments, the concept must be incorporated into human factors design efforts in each of these arenas. A theory of situation awareness that clearly defines the construct and its relation to human decision making and performance in dynamic systems is needed to fulfill this mission.

A Theory Of Situation Awareness

Although some have continued to argue that relatively little is known about SA, this belies the vast amount of work that has been done on the subject. While there is certainly a need for far more to be investigated relative to this construct, much can still be said about it. This information shall be presented in what Reason (1988) calls a framework model. That is, a model which is descriptive of the phenomenon and which synthesizes information from a variety of areas. Klein (1989b) states that a desired theory of situation awareness should explain: (a) dynamic goal selection, (b) attention to appropriate critical cues, (c) expectancies regarding future states of the situation and (d) the tie between situation awareness and typical actions.

Although this is certainly a tall order, much illumination on the subject can be created through an elaboration of the relationship between SA and various widely studied subjects of interest, including its relationship to information processing mechanisms, the impact of attention, workload and stressors, the role of expertise and automaticity, the influence of goal directed behavior in achieving SA, and the impact of SA on the decision making process. While this discussion probably cannot by any means be conclusive, it is intended to provide a common basis for discussion on SA and for much needed research on the topic.

Definitions, Descriptions and Models

Figure 1 provides a basis for discussing situation awareness in terms of its role in the overall decision making process. According to this model, a person's perception of the relevant elements in the environment, as determined from displays or directly by the senses, forms the basis for his/her situation awareness. Action selection and performance will proceed directly on the basis of that situation awareness which is the key step in the decision

process. Even the best trained decision makers will make the wrong decisions if they have inaccurate or incomplete situation awareness.

In order to proceed, a clear definition of situation awareness is needed first. While numerous definitions have been proposed (see Fracker, 1988), most are not applicable across different task domains. Referring to Figure 1, the following general definition of SA shall be used (Endsley, 1987a, 1988b): *"Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future."* To expand on this, the SA construct can be described in three hierarchical phases.

Level 1 SA – Perception of the elements in the environment. The first step is to perceive the status, attributes and dynamics of relevant elements in the environment. A pilot would perceive elements such as aircraft, mountains, or warning lights along with their relevant characteristics (e.g. color, size, speed, location). A tactical commander needs accurate data on the location, type, number, capabilities and dynamics of all enemy and friendly forces in a given area and their relationship to other points of reference. An FMS operator needs data on the status of machines, parts, flows and backlogs. An automobile driver needs to know where other vehicles and obstacles are, their dynamics, and the status and dynamics of one's own vehicle.

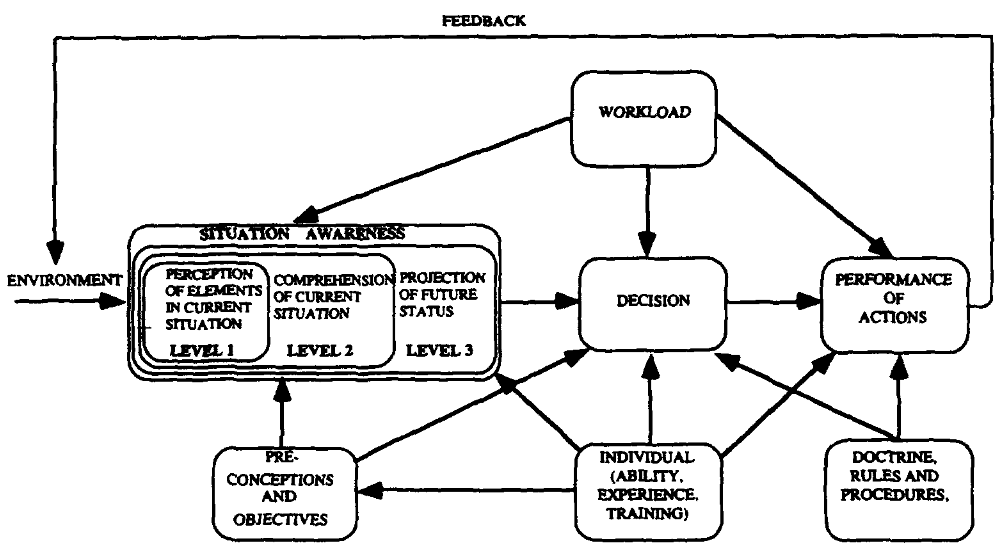


Figure 1. Situation awareness in dynamic decision making

Level 2 SA – Comprehension of the current situation. Comprehension of the situation is based on a synthesis of disjointed Level 1 elements. Level 2 SA goes beyond simply being aware of the elements which are present, to include an understanding of the significance of those elements in light of pertinent operator goals. Based upon knowledge of Level 1 elements, particularly when put together to form patterns with the other elements (gestalt), the

decision maker forms a holistic picture of the environment, comprehending the significance of objects and events. For example, a military pilot or tactical commander must comprehend that the appearance of three enemy aircraft within a certain proximity of each other and in a certain geographical location indicates certain things about their objectives. The operator of a complex power plant needs to put together disparate bits of data on individual system variables to determine how well different system components are functioning, deviations from expected values, and the specific locus of any deviant readings. In these environments, a novice operator might be capable of achieving the same Level 1 SA as more experienced decision makers, but may fall far short of being able to integrate various data elements along with pertinent goals in order to comprehend the situation as well.

Level 3 SA – Projection of future status. It is the ability to project the future actions of the elements in the environment, at least in the very near term, that forms the third and highest level of situation awareness. This is achieved through knowledge of the status and dynamics of the elements and a comprehension of the situation (both Level 1 and Level 2 SA). For example, knowing that a threat aircraft is currently offensive and is in a certain location allows a fighter pilot or military commander to project that the threat aircraft is likely to attack in a given manner. This gives him the knowledge (and time) necessary to decide on the most favorable course of action to meet his objectives. Similarly, an air traffic controller needs to put together various traffic patterns to determine which runways will be free and where potential collisions are. An automobile driver also needs to detect future possible collisions in order to act effectively, and an FMS operator needs to predict future bottlenecks and unused machines in order to schedule.

In addition, a crucial factor in understanding SA in a given environment will rest on a clear elucidation of the *elements* in this definition. These elements, however, will be specific to individual systems and contexts, and, as such, are the one part of SA that cannot be described in any valid way across arenas. Although each relies on SA, it simply is not realistic nor appropriate to expect the same elements to be relevant to both a pilot and to an FMS operator, for example. Nonetheless, these elements can be, and should be, specifically determined for various classes of systems. Endsley (in press) presents a methodology for accomplishing this and describes such a delineation for air-to-air fighter aircraft.

Several other aspects of situation awareness should be mentioned at this point. First, while SA has been discussed as a person's knowledge of the environment at a given point in time, it is highly temporal in nature. That is, SA is not always acquired instantaneously, but is built up over time. Thus, it takes into account the dynamics of the situation which are only acquirable over time. These dynamics are then used to project the state of the environment in the near future. So SA, while a model of the environment at any point in time, includes temporal aspects of that environment, relating to both the past and the future.

Secondly, it has been observed that SA is highly spatial in nature in many contexts. Pilots and air traffic controllers, for instance, are highly concerned with the spatial relationships between multiple dynamic aircraft. In many other fields, there may also be a concern for the spatial as well as functional relationships between system components. In any case, one important aspect that is often spatially determined is the specification of just which aspects of the environment are important for SA. Situation awareness can be conceived of as concerning only the subset of the environment which is considered relevant to the task(s) at hand. This boundary can be seen to shrink or expand as various tasks present themselves in a dynamic environment. The boundary may change temporally, spatially or functionally through refocusing on different system components within the problem space or by changing

the boundaries of the problem space itself. In a piloting context, for example, the boundary may shift spatially and temporally to include different aircraft depending on current goals and tasks. In other contexts, such as a manufacturing or power plant environment, the boundary may shift functionally to include different sub-systems. In both cases, a mental boundary exists, within which are elements relevant to SA. Within this boundary, the SA region may be further subdivided into levels of importance for SA or may assume a relevance continuum, depending on the problem context. Figure 2 shows an example of such boundaries for fighter pilot SA.

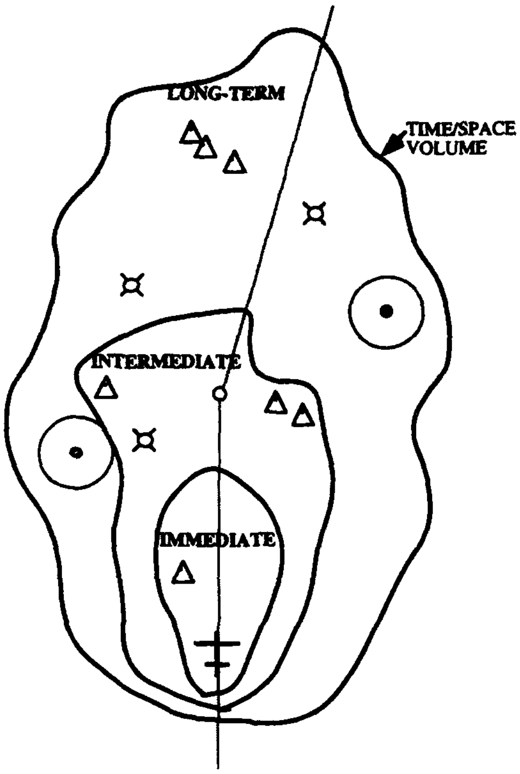


Figure 2. Zones of interest – Fighter pilot SA

As a side issue, it is also possible to talk about SA in terms of teams, as well as individuals. In many situations several individuals may work together as a team to make decisions and carry out actions. In this case, one can conceive of *overall team SA* where each team member will have a specific set of SA elements about which they are concerned, as determined by their responsibilities within the team. SA for an entire crew can be represented as shown in Figure 3. Some overlap between each team member's SA requirements will be present. It is this subset of information which constitutes much of inter-team coordination. That coordination may occur as a verbal exchange, as a duplication of displayed information, or by some other means. As such, the quality of team members' SA

of shared information may serve as an index of inter-team coordination and human-machine interface effectiveness.

Overall team SA can be conceived of as the degree to which every team member possess the SA required for his or her responsibilities. This is independent of any overlaps in SA requirements that may be present. If each of two team members needs to know a piece of information, it is not sufficient that one knows perfectly, but the other not at all. Each and every team member must have SA for all of his or her own requirements, or become the proverbial weak link in the chain.

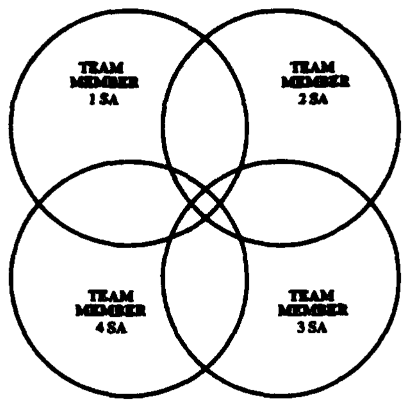


Figure 3. Team situation awareness

As expert operators of dynamic systems often act directly on the basis of their situation awareness, it will greatly behoove designers of these systems to maximize the situation awareness provided by a system. A situation awareness approach strives to present information which is integrated (Level 2 SA) and which provides support for the operator's projection needs (Level 3 SA) in order to best support the operator's goals. In order to effectively incorporate this concept into current research and design efforts, a more complete understanding of the mechanisms underlying the SA construct is needed.

Information Processing, Perception And Memory

Although little research has been directed at situation awareness specifically, a great deal of research in psychology has been devoted to more general aspects of human cognition. While much debate still continues in the psychology community as to the exact structure and nature of information processing mechanisms, a detailed discussion of various theories regarding each lies beyond the scope of this paper. Rather, the relationship between SA and these mechanisms, as generally understood, will be explored.

In combination, the mechanisms of short term sensory memory, perception, working memory and long term memory form the basic structures on which SA is based. Figure 4 shows a schematic description of the role of each of these structures in the SA process.

According to most research on information processing (for a review see Norman (1976) or Wickens (1984)), the environment is initially processed in parallel through preattentive

sensory stores where certain properties are detected, such as spatial proximity, color, simple properties of shapes, or movement (Neisser, 1967; Treisman and Paterson, 1984), providing cues for further focused attention. Those objects which are most salient, based on preattentively registered characteristics, will be further processed using focused attention to achieve perception. The deployment of attention in the perception process acts to present certain constraints on a person's ability to accurately perceive multiple items in parallel, and, as such, is a major limit on SA.

In addition to external factors, attention and perception can be directed by the contents of both working memory and long-term memory. Specifically, advance knowledge of the position of information (Posner, Nissen, and Ogden, 1978), the form of the information (Barber and Folkard, 1972), the spatial frequency (Davis, Kramer, and Graham, 1983), the color (Humphreys, 1981), or the overall familiarity and appropriateness (Biederman, Mezzanotte, Rabinowitz, Francolin, and Plude, 1981; Palmer, 1975) can each significantly facilitate perception. Information stored in long term memory also serves to shape the perception of objects in terms of known categories or mental representations. Ashby and Gott (1988) found that subjects based categorization upon integrated information about the object, typically in a deterministic, nearly optimal manner. Hinsley, Hayes and Simon (1977) found that this occurred almost immediately.

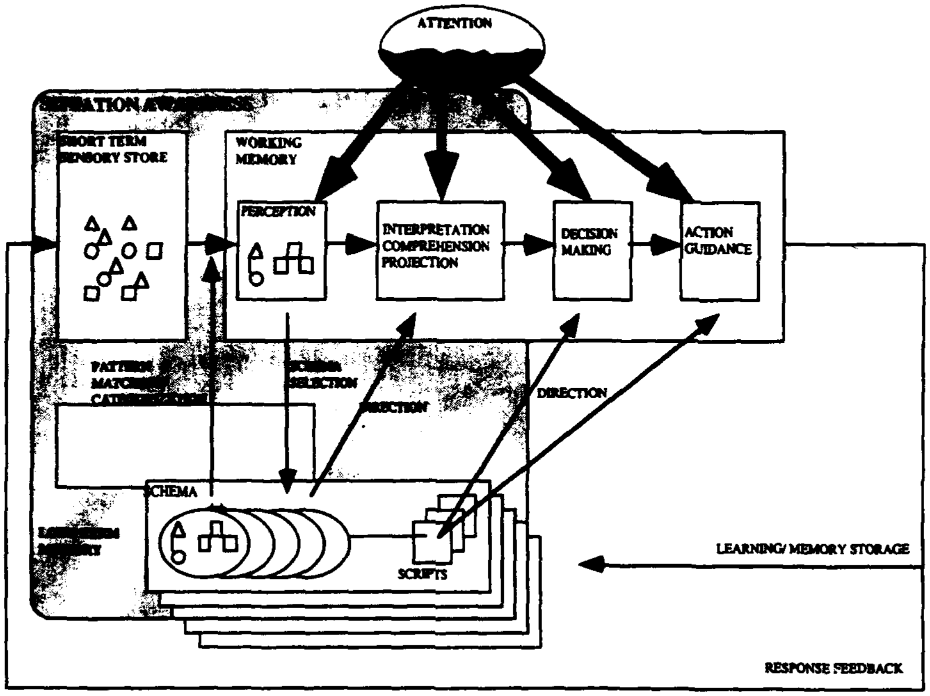


Figure 4. Mechanisms of situation awareness

The perception of the elements in the environment, the first level of SA, therefore, will be largely guided by the contents of working and long term memory stores to direct attention, recognition and categorization. The operator will direct his/her attention to look for what is

expected or needed based on memory stores, and will interpret what is perceived in light of them. Fracker (1989), for instance, in a study of pilot SA, showed that a limited supply of attention was allocated to environmental elements on the basis of their ability to contribute to task success. Because the supply of attention appears to be limited, improvements in SA on some elements may mean decrements in SA on others once the limit is reached. And this limit may occur rather quickly in complex environments.

Once perceived, information is stored in working memory, a limited capacity system for holding and manipulating information, as required. Working memory can be fed with information from either the environment or from long term memory storage. In the absence of other mechanisms, most of a person's active processing of information in the dynamic decision making environment must occur in working memory. New information must be combined with existing knowledge and a composite picture of the situation developed. Projections of future status and subsequent decisions as to appropriate courses of action will occur in working memory as well. In this circumstance, a heavy load is imposed on working memory, as it is taxed with simultaneously achieving the higher levels of situation awareness, formulating and selecting responses and carrying out subsequent actions. Wickens (1984), for instance, has stated that prediction of future states (the culmination of good SA) imposes a strong load on working memory by requiring the maintenance of present conditions, future conditions, rules used to generate the latter from the former, and actions that are appropriate to the future conditions. Fracker (1987) hypothesized that working memory constitutes the main bottleneck for situation awareness.

In actual practice, however, long term memory structures can be used to circumvent the limitations of working memory. The exact organization of knowledge in long-term memory has received diversified characterization. Typical organization schemes include (a) episodic memory, allowing for the storage of self-referent episodes, including temporal information and spatial relationships between perceived items (Tulving, 1972), (b) semantic networks, allowing for pieces of information to be represented based on their conceptual meaning and linked to each other based on certain relationships between items, including explicitly stated relations or rules (Tulving, 1972), and (c) schema, a general knowledge structure which serves to select and organize incoming information into an integrated meaningful framework including frames, prototypes and scripts (Mayer, 1983).

Schema can provide coherent frameworks of understanding, encompassing highly complex system components, states and functioning. Much of the details will be lost when information is coded in this manner, but the information will become more coherent and organized for storage, aiding retrieval and further processing. A single schemata may serve to organize several sets of information, and as such will have variables which can be filled in with the particulars for the particular case being considered. A script, a special type of schemata, provides sequences of appropriate actions for different types of task performance. Ties between schema and scripts can greatly facilitate the cognitive process, as an individual does not have to actively decide on appropriate actions at every turn, but will automatically know the actions to take for a given situation based on its associated script.

Another related form of memory organization which has been used frequently is *mental model*. Rouse and Morris (1985) define mental models as "mechanisms whereby humans are able to generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future states". Doing so depends on the person's ability to abstract and utilize cues from the environment. They state that experts will develop mental models in a shift from representational to abstract codes. From this definition, mental models can be described as complex schema that are used to model the

behavior of systems. Therefore, a mental model can viewed as a schemata for a certain system.

In addition, a second type of mental model will be considered. A *situational model* (or situation model), a term used by VanDijk and Kintsch (1983) and by Roschelle and Greeno (1987), will be defined as a schemata depicting the current state of the system model (and often developed in light of the system model). Rasmussen (1986) also used the term "internal dynamic world model" with the same general meaning. The terms situation model and situation awareness will be defined here as equivalent. A situation model can be matched to schema in memory depicting prototypical situations or states of the system model in order to activate associated goals or scripts. This process will be described in more detail.

Schema and mental models are developed as a function of experience with a given environment. In the beginning, a person who is a novice in an area may have only a vague idea of system components and sketchy rules or heuristics for determining the behavior they should employ with the system. With experience, recurrent situational components will be noticed along with recurrent associations and causal relationships. This forms the basis for early schema or model development. Holland, Holyoak, Nisbett and Thagard (1986) provide a thorough description of the development of mental models. According to their description, an individual will learn: (a) categorization functions (P) that allow people to map from objects in the real world to a *homomorphism* (many-to-one mapping), i.e., their mental model, and (b) model transition functions (T') that describe how objects in the model will change over time based on transition functions in the real world (T). By comparing the predictions of their internal model at some time, t2, with the actual states of the system at t2 through repeated experience, individuals will progressively refine their models. They may develop more specific and more numerous categorization functions to allow for more accurate predictions based on more specific object characteristics, or they may develop better transition functions. Holland et. al. 's explanation also includes a *Q-morphism* in which default expectations for the system are provided in a higher layer of the model. These default values may be used by individuals to predict system performance unless some specific exception is triggered, in which case the appropriate transition function for that classification will be used.

This description provides for the development of complex mental models through experience. The main key to using these models rests on the ability of the individual to recognize key features in the environment that will map to key features in the model (through P). The model then provides for determining associations between components and predictions of the behavior and status of the system over time. These structures can provide for much of the higher levels of SA (comprehension and projection) without loading working memory. Where scripts have been developed for given situation conditions (through experience in much the same way), much of the load on working memory for generating alternate behaviors and selecting among them is also diminished. These mechanisms allow a person to simply execute a predetermined action for a given recognized class of situations (based on the situation model or SA). And the current situation need not be exactly like one encountered before due to the use of default values in the Q-morphism and the use of categorization mapping. As long as a mapping can be made into relevant categories, a situation can be recognized, comprehended in terms of the model, predictions made and appropriate actions selected. Of prime importance is that this process can be almost instantaneous due to the superior abilities of human pattern matching mechanisms.

For novices or those dealing with novel situations, decision making in dynamic environments is an arduous task, requiring detailed mental calculations based on rules or heuristics, placing a heavy burden on working memory. Where experience has allowed the

development of long-term memory structures, pattern matching between the perceived elements in the environment and existing schema and mental models will occur on the basis of pertinent cues. When these long term memory structures exist, they can be utilized to provide the required comprehension and future projection required for the higher levels of SA, thus off-loading working memory requirements substantially. When scripts have been developed, tied to these schema, the entire decision making process will be greatly simplified.

An important aspect of situation awareness that should be mentioned at this point concerns a person's confidence level regarding that SA. Certainly, it has been widely discussed that a person can have a confidence level associated with some decision. A person may also have a confidence level associated with information that has been acquired to make that decision, based on the reliability or source of the information. The confidence level associated with information can impact the decisions that are made using that information (Norman, 1983). An important aspect of SA, therefore, is the person's confidence concerning that SA.

Holland, et. al. (1986) consider this in their discussion of mental models. They hypothesize that there is a degree of uncertainty associated with the model transition function (T') that will provide for confidence levels associated with predictions from the model. Similarly, one could hypothesize a degree of uncertainty associated with (P), the mapping from the real world to the internal model, in that there may be uncertainty about the validity of features used to make that mapping. For example, if three sources of information indicate a certain object is an apple, but one source indicates it is an orange, the object may be characterized in the internal model as an apple, but with an uncertainty factor attached to it.

VanDijk and Kintsch (1983), in their work on speech understanding, have conceptualized a *context model* as an underlying factor of this process which represents facts about a piece of information's source, including intentions, speaker, the context of the speech, etc. This context model can be instrumental in a person's ability to use conflicting information from varied sources with varying reliability, and perhaps questionable motives, by taking these factors into account in the decision process as well as the stated facts. If this concept is borrowed and applied to a more general class of models, any given situation model may include uncertainty regarding (a) the accuracy of the components of the model in the form of an associated context model and correspondingly the mapping of world information to the internal model (P), and (b) the accuracy of predictions from the model based on uncertainty surrounding the transition function (T'). An important outcome of these confidence levels is that they still allow people to make decisions effectively, despite numerous uncertainties, yet small shifts in factors underlying the uncertainties can dramatically change resultant conclusions.

Attention, Workload and Stress

Attention serves as an important constraint on situation awareness. Direct attention is needed for not only perception and working memory processing, but also for decision making and forming response executions. In complex dynamic environments, attentional demands due to informational overload, complex decision making and multiple tasks can quickly exceed limited attention resource capacities. A procedure of successive environmental sampling is often employed as a means of coping with the problem, where sampling strategies are guided by internal system models. Typically, humans will have several failings in this sampling process, including non-optimal strategies based on misperceptions of statistical

properties of elements in the environment and limitations of human memory (Wickens, 1984).

The phenomenon of visual dominance, a bias towards information that is presented in the visual modality, acts to further limit human attentional capacities (Posner, Nissen, and Klein, 1976). Kahneman (1973) has proposed that attentional resources can be increased to some degree by physiological arousal mechanisms. Further relief to the limitations dictated by limited attentional resources can be met through the capability of people to divide their attention under certain circumstances. Damos and Wickens (1980) have found that sharing of attentional resources is a skill which can be learned and at which some people excel over others. Thirdly, limitations of attention can be circumvented through the development of automaticity, which will be discussed in more detail later.

In addition to attention limitations, several types of stress factors exist which may also act to impact situation awareness and dynamic decision making. These factors include (a) Physical stressors – noise, vibration, heat/cold, lighting, atmospheric conditions, drugs, boredom or fatigue, cyclical changes, and (b) Social/Psychological stressors – fear or anxiety, uncertainty, importance or consequences of events, aspects of task effecting monetary gain, self-esteem, prestige, job advancement or loss, mental load, and time pressure (Hockey, 1986; Sharit and Salvendy, 1982). In many dynamic systems, high mental workload is often a prime stressor of particular importance.

Mandler (1982) states that these stressors "are effective to the extent that they are perceived as dangerous or threatening". That is, they are only stressors if the person perceives them as being stressing. A large interpretive component exists in the process. A certain amount of stress can act to actually improve performance by increasing attention to important aspects of the situation. A higher amount of stress can have extremely negative consequences, however, as accompanying increases in autonomic functioning and aspects of the stressors can act to demand a portion of a person's limited attentional capacity (Hockey, 1986).

Stressors can effect situation awareness in a number of different ways. The first and probably most widespread finding is that under various forms of stress, people tend to narrow their field of attention to include only a limited number of central aspects (Bacon, 1974; Baddeley, 1972; Bartlett, 1943; Callaway and Dembo, 1958; Davis, 1948; Eysenck, 1982; Hockey, 1970). A decrease in attention is generally observed for peripheral information, those aspects which attract less attentional focus, under perceived danger (Bacon, 1974; Weltman, Smith, and Egstrom, 1971). Broadbent (1971) found that there was an increased tendency to sample dominant or probable sources of information. Sheridan (1981) has termed this effect "cognitive tunnel vision". This is a critical problem for situation awareness, leading to the neglect of certain features in favor of others. In many cases, such as in emergency conditions, it is those factors outside the operator's perceived central task that prove to be lethal.

Premature closure, arriving at a decision without exploring all information available, has been found to be more likely under stress (Janis, 1982; Keinan, 1987; Keinan and Friedland, 1987). Janis (1982) and Wright (1974) both found that less information was considered under stress, and Wright (1974) found that subjects under time pressure attended more to negative information. Woodhead (1964) found that performance decrements during intermittent noise stress in a calculation task occurred only during the information input stage.

In addition, several authors have found that scanning of stimuli under stress is scattered and poorly organized (Keinan, 1987; Keinan and Friedland, 1987; Wachtel, 1967). Complex tasks with multiple input sources are particularly sensitive to the effects of stressors

(Broadbent, 1954; Jerison, 1957; Jerison, 1959). It would seem then that stress significantly affects the early stage of the decision making process that is involved in the recognition and assessment of the situation at hand. It is expected that there will be a significant impact of stress on situation awareness on this basis.

A second way in which stress may impact situation awareness is through working memory. Working memory is in high demand during many phases of the decision making process, when novel stimuli must be interpreted and comprehended, a prediction of future states determined, and appropriate actions generated (Wickens, 1984). Many authors have found significant decrements in working memory capacity and retrieval during noise stress and anxiety (Hockey, 1986; Mandler, 1979). Wickens, Stokes, Barnett and Hyman (1988) found that optimality of performance was negatively affected by stress only on decision tasks with a high spatial component, however, and not on those with purely a high working memory or long-term memory component.

The exact impact of these effects on situation awareness and decision making will be varied. In tasks with a high working memory load, such as those requiring mulling of alternatives or generation of novel actions, a significant impact would be expected. As a great deal of expert decision making may utilize long-term memory structures in a pattern-matching process, however, the effect may be minimal in those cases. Endsley (1989b) found that situation awareness information was accessible from long-term memory in expert fighter pilots, supporting this view. Stress, therefore, may impact situation awareness and dynamic decision making through working memory restrictions only in some cases.

While there have been some findings associated with more general impacts of stress during decision making, such as the increased use of over-simplified decision rules (Janis, 1982) and over-confidence in decisions (Broadbent and Gregory, 1965; Sieber, 1974), the majority of research in the area points to two specific areas of impact (a) the information input or situation awareness stage, and (b) information processing required in working memory when other mechanisms are not available.

Expertise and Automaticity

The role of expertise in a given task environment has already been discussed extensively. As an individual acquires experience in an area, a mental model may be developed to aid in the higher levels of SA and in action selection. Without such a model, only loose rules and heuristics will be available to guide decision making, resulting in a limited capacity to act appropriately in a wide variety of situations and a heavy load on working memory. This effectively acts to increase the decision time required and to reduce the ability to deal with very complex systems. With increasing expertise, as Dreyfus (1981) points out, the individual will include in decision activities a consideration of situational components, a recognition of component salience, a holistic style of situation recognition and an intuitive decision style.

In developing expertise, a form of automaticity can be acquired. Logan (1988) provides a detailed discussion of automaticity which he maintains allows a direct-access, single-step retrieval of actions to be performed from memory. Automatic processing tends to be fast, autonomous, effortless and unavailable to conscious awareness in that it can occur without attention. This automaticity can result through (a) a reduction in the resources demanded for the task, or (b), for which Logan argues, a direct retrieval of information from memory,

novice performance therefore being limited by a lack of knowledge rather than a lack of resources.

Automaticity as related to cognitive processes in dynamic decision making can take two forms. In the first, there is automaticity of information retrieval from memory on the basis of situation awareness, using the schema mechanisms described. In this process "attention to an object is sufficient to cause retrieval of whatever information has been associated with it in the past" (Logan, 1988, p. 587). In this circumstance, the individual is conscious of the situational elements which triggered the automatic retrieval of information from memory (SA), but probably will not be conscious of the mechanisms used for arriving at the resultant action selection. That is, the individual knows the *what,* but not the *how,* as expressed by Dreyfus (1981). If asked to explain why a particular decision was made, an individual will usually have to construct a rationale using logical processes to provide an explanation of the action they actually chose in an automatic, non-analytic, manner. This manner of automaticity is very typical of expert decision makers.

A second form of automaticity speculates on processing without conscious situation awareness. Evidence for this phenomenon is rather sketchy. Nisbett and Wilson (1977) describe many studies in which people can have affective processes (emotions, opinions, attitudes, perceptions) altered by some stimuli without a reportable awareness that these factors have even changed or that the stimuli were involved in the change. This however, does not mean that they were unaware of the stimuli, only that they could not verbalize or were unaware of the *causal link* between the stimuli and the resultant change. Often when pressed for an explanation of their behavior, subjects constructed plausible and/or socially acceptable causal explanations, most likely using available mental models, which may or may not have involved the critical stimuli.

In relation to typical decision making or problem solving situations, Nisbett and Wilson report several cases in which the solution to some problem occurs either (a) without conscious attention to the problem (attention being directed at other problems or activities), or (b) without the ability to report on the critical stimuli leading to the solution. In the later case, it is again the causal link that they cannot reliably identify, rather than the existence of the stimuli. In the first case, it is likely that the cognitive processes leading to the eventual solution did indeed occur below the threshold level of conscious awareness. In these cases, however, all of the requisite situational information was already stored in memory. Problem solution most likely occurred either through (a) the development of a better model using synthesis and revision of existing models, or (b) the adoption of a whole new model for problem solving, a process which may have occurred through some internal process or may have been triggered by some external stimuli, which they again may or may not be able to report as causal. In all of these cases it would appear that again the *how* becomes occluded through the use of automatic processes, but the *what* is still available to awareness. The one exception to this statement is subliminal stimuli, which have been shown to modify affective processes. Evidence for the role of subliminal stimuli on typical dynamic decision making, as opposed to affective processes, is less apparent.

Reason (1984) argues that some very low level of attention is required at a minimum for all activity, even automatic processes, in order to bring appropriate schema into play at the right times and to restrain unwanted schema from jumping in. At this level of attention, there would be no awareness (equated with consciousness) of the detailed procedures. Once a plan has been put into motion, it serves to execute scripts and process schema as instructed. An extreme example of the possibility of decision making without conscious SA is that of the individual driving home from work who follows the same predetermined path, stops at stop

lights, responds to brake lights and goes with the flow of traffic, yet can report almost no recollection of the trip. Did this person truly operate with no conscious awareness? Or, is it rather that only a low level of attention was allocated to this routine task, keying on critical environmental features which automatically evoked appropriate actions? The low level of consciousness simply did not provide sufficient salience to allow that particular drive home to be retrieved from memory as distinguishable from a hundred other such trips. I would argue, in agreement with Reason, that this later alternative is far more likely.

So how much situation awareness do people who are operating at this level of automaticity have? And how much do they need to function effectively? These are very good questions, for which more data are needed. At this point, it can only be said that several authors (Jacoby and Dallas, 1981; Kellog, 1980; Tulving, 1985) have found that even when effortful processing is not used, information can be retained in long-term memory and can effect subject responses. The major implications of the use of automatic processes are (a) good performance with minimal attention allocation, (b) significant difficulty in accurately reporting on the internal models used for such processing and possibly on reporting which key environmental features were related, and (c) unreliability and inaccuracy of reporting on processes after the fact. The primary hazard created by automatic cognitive processing is the increased risk of being less responsive to new stimuli, as these processes operate with limited use of feedback. That is, a lower level of SA could result, decreasing decision timeliness and effectiveness in non-typical situations. Again, more data is needed on this subject.

Goals, SA and Dynamic Decision Making

Situation awareness is not generally thought of as a construct which exists solely for its own sake. In general, SA is important as a part of decision making regarding some system. As such, it is integrally linked with both the context and the decisions for which the SA is being sought. In light of this, dynamic decision making will be discussed in so far as it impacts on SA and SA impacts on it. This discussion will be broken into (a) the role of goal directed decision making in acquiring SA, and (b) the impact of SA on the decision process itself.

Goals form the basis for most decision making in dynamic environments. Several postulates serve as the basis for this assertion.

1. People operate on the basis of goals or objectives. While some behavior may be random and spontaneous, it can be stated that people often operate in light of goals (to be happy, to graduate from school, to get a raise, to win the battle, etc.). Furthermore, they may have more than one goal operating simultaneously, and these goals may conflict (e.g. "stay alive" and "kill enemies").

2. People are active participants in their environment. They may actively seek to alter the environment in order to meet their goals and objectives.

3. People require knowledge of their environment in order to effectively change it. In order to alter the environment in a way that is consistent with their goals, an assessment of the state of the environment is necessary. Other than through blind luck, this is generally required. Furthermore, it can be stated that people are not simply helpless

recipients of data from the environment, but may actively seek to search out data in light of their goals.

4. Decision making can be seen as a process of selecting activities that will alter the environment to either directly or indirectly meet goals. Activities may be selected that directly meet a goal (e.g., a shot is fired, downing an enemy aircraft) or which are indirectly in accordance with meeting a goal (e.g., the airplane in maneuvered into a good firing position).

Based on these basic premises, dynamic decision making is characterized as including goals, the plans selected for achieving those goals, and activities for carrying out the plans. Associated with each goal is a projected state reflecting what the world will be like if that goal is reached. Goals will be selected whose projected state is desirable to the decision maker. Similarly, associated with each plan is a projected state reflecting what the world will be like if the plan is carried out. In general, plans will be selected whose projected states match the goals. Specific activities are selected that are in accordance with the selected plans.

This description is very similar to the Image Theory proposed by Beach and Mitchell (1987). According to Image Theory, the only decisions made are (a) adoption decisions – the selection of compatible goals, plans and activities, and (b) process decisions – decisions regarding the continuance of already active goals, plans and activities, based upon their compatibility. Image Theory, however, neglects an important component of dynamic decision making – the person's assessment of the state of the environment in light of his/her goals and plans. A major activity in dynamic decision making involves the gathering of information to form an internal model of the environment. This internal model, the person's situation awareness, is a primary input to the decision making process.

In what Casson (1983) has termed a top-down decision making process, the person's goals and plans will direct which aspects of the environment are attended to in the development of SA. Activities will then be selected by the decision maker which will bring the perceived environment into line with the person's plans and goals. Conversely, in a bottom-up process, patterns in the environment may be recognized which will indicate to the decision maker that different plans will be necessary to meet the goals or that different goals should be activated. It is this process of assessing the environment and acting upon it that makes for dynamic decision making. The decisions that occur are not only in regard to the selection and continuance of goals and plans, but also concern the selection of activities that will bring the perceived situation state into alignment with the desired state.

How does this relate to the important role of schema in dynamic decision making? The model in Figure 5 can be used to visualize this relationship. Mental models of systems can be seen to exist as set (although slowly evolving) memory structures. Independently, individuals may form a set of goals that relate to some system. These goals can be thought of as ideal states of the system that they wish to achieve. The same set of goals may exist frequently for a given system, or they may change often. Conversely, a set of goals may relate to more than one system model. A person's current goal(s) (selected as the most important among competing goals) will act to direct the selection of a mental model to aid in devising a strategy for achieving the goal. It will also determine the frame (Casson, 1983), or focus, on the model that is adopted. Plans are then devised for reaching the given goal, using the projection capabilities of the model. A plan will be selected whose projected state best matches the goal state.

Associated with the system model are also various scripts for achieving specific outcomes. Schank and Abelson (1977) explain the relationship between plans and scripts, stating that plans form the general mechanisms underlying scripts. That is, a plan is basically a serious of actions devised to reach a goal. A script is one means of realizing a plan. When scripts are available for executing the selected plan, they will be employed. When scripts are not available, actions will have to be devised which will allow for plan completion. Again, the projection capabilities of the system model will be used to accomplish this.

Once determined, the individual carries out the actions, which in turn impacts on the state of the environment. In addition, the environment may change over time of its own volition, either through naturally occurring events or through the actions of others. As an ongoing process, therefore, the individual must observe the current state of the environment. The individual's attention to environmental features will be directed by the activated model and interpreted in light of it. This forms the basis for the individual's expectations concerning what will happen. The model that is active provides a future projection of the status of key environmental features. Over time, the individual will be matching the observed situation to an internally held projection of system states. This provides expectations for not only what will be observed, but also for what should not be observed. When the two models (observed and internal projection for that time) match, all is well. When they do not match because values of some parameter are different, an event occurs that should not, or an event does not occur that should, this signals the individual that something is amiss, and indicates a need for a change in goals or plans due to a shift in situations, a revision of the system model, or selection of a new system model.

In this way, new relevant schema (models, frames, or prototypical situation patterns) may be evoked by pattern-matching critical features of the environment to competing internal schema. The overall decision making process can be viewed, therefore, as a dual process whereby active schema or mental models are dictating which information to focus attention on (conceptually driven), and simultaneously the presence of certain objects or attributes in the environment will activate new schema in memory (data driven).

This process can act to change current goal selection by altering the relative importance of goals, as each goal can have antecedent rules governing situations in which each needs to be invoked over the others. Wherein multiple goals are compatible with each other, several may be active at once. Where goals are incompatible, their associated priority level for the situation determines which shall be invoked. Similarly, plans may be altered or new plans selected if the feedback provided indicates that the plan is not in accordance with its projections, or when new goals require new plans. Through learning, these processes can also serve to create better system models, allowing for better projections in the future.

So what will happen if no mental model exists? The individual will be forced to (a) act randomly to accomplish goals, (b) build a new model through repeated experience with the environment, (c) adapt a model of a similar system to attempt some means of guiding action selection, or (d) some combination of these tactics. In any case, it is to be expected that this individual's selection of plans, activities and often even sub-goals will be highly error prone and sub-optimal for overall goal completion, unless the adapted model is a very good match to the new system or the person is extremely lucky.

To give an example of this process, in a military aircraft environment, a pilot may have various goals such as stay alive, kill enemy aircraft, and bomb a given target. General goals may have more specific sub-goals such as navigate to the target, avoid detection, avoid missiles, employ missiles, etc. The pilot would choose between goals and sub-goals based on their relative importance and the situation. Staying alive is a priority goal for example, which generally is active (except in extreme kamikaze circumstances). A pilot may alternate

between the goals of bombing a target and killing an enemy aircraft based on the predetermined criticality of each goal's success to the overall war and the specifics of the situation, such as the likelihood of each goal's success based on current system parameters and the current distance to each.

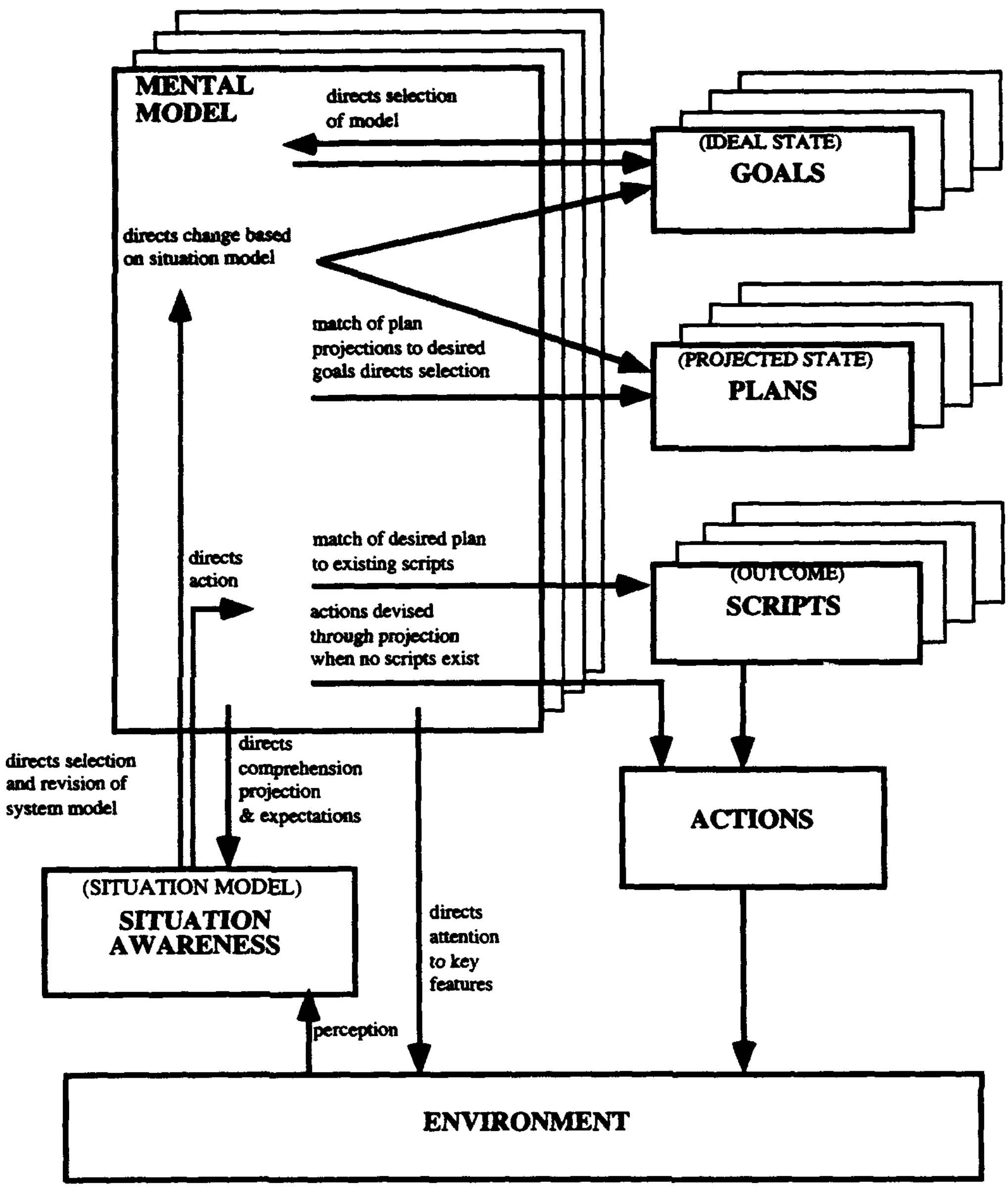


Figure 5. Role of goals and schema in SA and dynamic decision making

The current goal would indicate the model and frame to be active. A model for missile employment might direct attention towards key features such as dynamic relative positions of own and threat aircraft (location, altitude, airspeed, heading, flight path) and current weapon selection including weapon envelope/capabilities, current probability of kill, and rate of change of probability of kill. If this model was active, the pilot would be inclined to seek out and process those key elements of the environment. Use of the resultant situation model, in conjunction with the system model, would allow him to determine the optimal manner of employing his aircraft relative to the enemy aircraft and missile launch timing (plans and actions).

While carrying out this goal, however, the pilot will also be alert to critical features which might indicate that a new model should be activated. If the pilot detected a new threat, for example, his goals might change so that he would cease to operate on the missile employment model and a threat assessment model would be activated. The model selected, when detailed enough, can be used to direct situation comprehension, future projection, and decision making. A threat assessment model might include information as to what patterns of threats and threat movements constitute offensive versus defensive activities, for example. Future threat movements might be predictable from the model through a classification of current threat movements into known tactics. Appropriate tactics for countering given threat actions might also be resident in the form of scripts, greatly simplifying decision making. Much of an expert's situation awareness may draw on these types of structures, as they provide the necessary framework with which to quickly organize, comprehend and predict complex system behavior.

Situation Awareness and the Decision Making Process

In addition to forming the basis for decision making as a major input, situation awareness may also impact the process of decision making itself. There is considerable evidence that a person's manner of characterizing a situation will determine the decision process chosen to solve a problem. Manktelow and Jones (1987) review the literature concerning deductive problem solving. They show, through numerous studies, that the situational parameters, or context, of a problem largely determine the ability of individuals to adopt an effective problem solving strategy. It is the situational specifics which determine the adoption of an appropriate mental model, leading to the selection of problem solving strategies. In the absence of an appropriate model, people will often fail in solving a problem correctly, even though the same logical process need be employed as in a problem they are familiar with.

Other evidence suggests that even the way a given problem is presented can determine how the problem is solved (Bettman and Kakkar, 1977; Herstein, 1981; Sundstrom, 1987; Tversky and Kahneman, 1981). The most simple explanation for this is that different problem framings can induce different information integration (situation comprehension). And, as has already been discussed, it is a person's situation comprehension as a whole that determines selection of a mental model to use for solving the problem. Thus, it is not only the detailed situational information (level 1 SA), but also the way the pieces are put together (level 2 SA) that directs decision strategy selection.

But when will people use a situationally based decision process, as described by Dreyfus (1981) and Klein (1989b), and when will they use other approaches? Many authors, including Hamm (1988), and Hammond (1986, 1988) have sought to classify human decisions into a continuum ranging from analytic to intuitive approaches. Hammond (1986),

for instance, proposes that task features such as complexity of the task structure, ambiguity and form of presentation determine whether a person will select an analytic or intuitive decision style for a particular task. Hamm (1988) augments this work by his finding that decision styles can shift between analytic and intuitive many times within a single problem and his recognition of the role of individual differences in decision style selection.

Both of these approaches, however, are limited due to their tendency to group all non-analytic styles into a single category called *intuitive*. This classification is highly misleading in that it treats decisions that are based on guessing or loose heuristics, indicating a somewhat arbitrary decision style of low expertise and effort, the same as those that use holistic situation recognition to draw upon a high-level of expertise for problem resolution. Truly, neither style is analytic. However, to call both intuitive, a term which Hammond (1986) has noted is generally considered inferior, biased and hazardous, does serious disservice to the effectiveness of non-analytic expert decision making.

To overcome this weakness, the continuum might more appropriately be represented as in Figure 6, on a scale ranging from arbitrary to analytic to holistic. Such a scale would more accurately reflect the expertise used in making decisions. Whereas, both arbitrary and holistic styles may be preferred over analytic styles by decision makers in situations involving time pressure, a holistic style can only be used when the decision maker has a sufficiently developed knowledge base. In situations without time pressure, any style could be used. Task factors, such as those noted by Hammond (1986) may induce a more analytic (as opposed to arbitrary style) on most decision makers, however those with sufficient expertise may still opt for a holistic style when possible due to its efficiency in resource utilization.

A useful point regarding this classification is that it is a continuum. That is, even within a single problem, a holistic style may only be used for those parts of the problem for which sufficient knowledge bases exist. Other parts of the problem may be solved analytically if rules and process are known, or arbitrarily if the rules are not known or if not deemed important enough to merit the extra effort required by analytic processes. (See Christensen-Szalanski (1978,1980) or Shugan (1980) for a discussion of the effects of time constraints and cost of thinking on selection of decision strategies.)

Errors In Situation Awareness

It is not the intention here to discuss all types of human error, for which several taxonomies exist (see Norman (1983), Rasmussen (1986) or Reason (1987) for a full discussion), but rather to investigate the factors that can lead to break-downs in the situation awareness portion of the dynamic decision making process. These break-downs can occur due to either incomplete or inaccurate situation awareness. The discussion will be, somewhat artificially, separated into those factors affecting SA at each of its three levels.

Level 1 SA. At the very lowest level, a person may simply fail to perceive certain information that is important for SA in the assigned task (incomplete SA). In the most simple case, this may be due to a lack of detectability or discriminability of the physical characteristics of the signal in question due to some physical obstruction preventing perception (visual barrier, auditory masking, etc.), or due to a failure of the system design to

make the information available to the operator. In extreme cases, the only cue a person will have regarding the presence of certain information will coincide with the occurrence of an error. Rasmussen (1986) gives the example of a person not realizing it's icy until he slips. In this case, the condition could only be discerned in conjunction with the error and not sufficiently in advance to allow for behavior modification to prevent the error. In other cases, due to luck, no error may result from the lack of SA, however, the potential for error would rise significantly.

Figure 6. Continuum of decision styles

In many cases where SA is incomplete, the relevant signals or cues are readily discernible, yet not properly perceived by the subject. There can be several underlying causes for this. In many complex decision making environments, there is an overabundance of information to take in. The real challenge is to simultaneously obtain an accurate reading on all relevant variables, which may be changing rapidly and/or physically separated. In such cases, the most frequent human adaptation is to employ successive data sampling in order to maintain a fair degree of accuracy on each of the relevant variables (Wickens, 1984). In this case, errors in SA would be small (determined by the amount of change in each variable between successive samples) and distributed across the various variables of concern.

Failures in the sampling process are commonplace, however. Under normal conditions, a failure in the sampling process may result from the lack of an adequate sampling strategy or internal model for directing sampling to relevant cues. Training or experience is needed to establish an effective sampling strategy for a task. Wickens (1984) has also noted that humans have several failings in the sampling process, including misperception of the statistical properties of elements in the environment and limitations of human memory (forgetting what has already been sampled). The phenomenon of visual dominance, a bias towards information that is presented in the visual modality, can act to further limit human sampling abilities (Posner, et al., 1976). The problem is exacerbated by the fact that it takes approximately 100 msec longer to switch attention between visual and auditory modalities than within a modality (LaBerge, VanGelder, and Yellot, 1971). In situations where both auditory and visual information is being delivered at the same rate (i.e. one is not more novel than the other) the auditory information will be less likely to be processed (Posner, et al., 1976).

Furthermore, some people appear to be better at dividing their attention across different tasks than others (Damos and Wickens, 1980). Martin and Jones (1984) have found cognitive errors to be significantly correlated with capabilities in distributing attention across tasks. So, while environmental sampling can be an effective means of coping with excessive SA demands, human limitations in sampling, attention and attention sharing can lead to significant Level 1 SA errors.

This problem is only compounded by the addition of stress. As already noted, stress can seriously impact the information input stage through premature closure, a switch in factors attended to, and deterioration of the scanning process. The narrowing of attention brought on by stress or high workload can lead to a total lack of SA on all but the factor being

concentrated on. In 1972 an L-1011 commercial airliner went down in the Florida Everglades because all of the crew members were so focused on a problem with the nose gear indicator that they failed to notice that the aircraft was descending, leading to a loss of the aircraft and 99 lives aboard (National Transportation Safety Board, 1973). This example is only one of many involving fatal consequences from attentional narrowing. A major problem with attentional narrowing is that often a person will be *sure* he or she is attending to the most important information, but there is no way to know that assumption is valid without knowing the values of the other variables. In other cases, the normal sampling strategy has merely been interrupted and not re-activated in a timely manner. In either case, attentional narrowing can lead to serious errors in SA.

Inaccurate SA, the belief that the value of some variable is different than it actually is, can also occur. In relation to Level 1 SA, this would occur through the misperception of a signal. For instance, seeing a blue light as green due to ambient lighting or seeing a 3 as an 8 on a dial.

Level 2 SA. Very often, however, inaccurate SA will be the result of an inability to properly integrate or comprehend the meaning of perceived data in light of operator goals. This can occur for several reasons. A novice will not have the mental models necessary for properly comprehending and integrating all of the incoming data or for determining which cues are actually salient to established goals. In the absence of a good internal model, one must either (a) accept low SA and thus be compromised in decision making, or (b) develop a new model or adapt an existing model to the task at hand. In this case, SA errors will exist in the form of incorrect or incomplete SA where the adapted or newly developed model fails to match to the new environment.

In other cases, a person may incorrectly select the wrong model from memory based on a subset of situational cues and use this model to interpret all data that is perceived. Mosier and Chidester (1991) found evidence that aircrews made "recognitional, almost reflexive judgment, based upon a few, critical items of information; and then spent additional time and effort verifying its correctness through continued situational investigation." This strategy can be effective. Mosier and Chidester found that the best performing crews obtained a substantial portion of their information after making a decision.

If the wrong mental model is selected initially based on a subset of cues, however, a *representational error* may occur. These errors can be a particularly troublesome, as pointed out by Carmino, Idee, Larchier Boulanger and Morlat (1988), in that it can be very difficult to realize the wrong model is active, since all new data are interpreted in light of it and possibly also due to confirmation bias, as discussed by Fracker (1988). Thus, data which should indicate one thing are actually taken to mean something quite different based on the incorrect model. Fracker also points out that an incorrect model may be selected due to the human biases of representativeness and availability.

Even when the correct model has been selected with which to interpret and integrate environmental stimuli, errors can occur. Certain pieces of data may be mismatched to the model or not matched at all, resulting in a failure to recognize a prototypical situation (Klein, 1989b; Manktelow and Jones, 1987). This may be due to problems of attentional limitations or due to some incompleteness on the part of the model.

In addition, an SA error could occur due to an over-reliance on default values embedded in a model (Manktelow and Jones, 1987). In general, when new situations are encountered where the known default values are not appropriate, the model is modified to include the new class of situations. Before this occurs, however, or if cues received have not flagged the

different situation, significant SA errors can occur by incorrectly assuming defaults for some variables. When no model exists at all, Level 2 SA must be developed in working memory. An inability to accurately perform this step in a timely manner, due to insufficient knowledge and limitations of working memory, particularly under stress, can also lead to inaccurate or unacceptable SA.

Level 3 SA. Finally, Level 3 SA may be lacking or incorrect. Even if a situation is clearly understood, it may be difficult to accurately project future dynamics without a highly developed mental model. Klein (1989b) has noted that some people are simply not good at mental simulation. The lack of a good model or attentional and memory limitations could account for this.

A few general underlying factors may also lead to SA problems. Martin and Jones (1984) have pointed out that people who have trouble with distributed attention may be having trouble in maintaining multiple goals. This could lead to considerable SA problems in complex systems, where the ability to juggle goals on the basis on incoming information is a necessity. An inability to keep multiple goals in mind could seriously degrade an operator's receptivity to highly pertinent data related to the neglected goal, leading to significant errors.

A second major type of error impacting SA relates to the role of habitual schema. Humans are creatures of habit. In the normal course of events, habitual schema will be automatically activated based on the presence of environmental cues. While this schema is active, the environmental cues will be processed in a predetermined manner. When a change needs to be made, however, problems can occur. A person leaving work and getting into their car may automatically embark on the "drive home" schema. If on a particular day the person wishes to stop at the store, he or she must change or interrupt the schema. Often, however, the person may arrive home to realize he or she completely forgot to make the desired detour.

While this has been termed a "slip of action" (Reason, 1984), it can also be shown to be related to SA. Under normal circumstances, environmental cues (the store sign) will be processed in light of current goals (stop at the store). While habitual schema are operating, however, the new non-habitual goal is suppressed, and seeing the store sign does not conjure up the associated goal of stopping. While the habitual schema is operating, the person either (a) is not receptive to the non-habitual cues, or (b) does not generate the appropriate higher level SA from the perception of the cues because the appropriate schema are suppressed.

The real question is how does a person know his or her SA is off-the-mark? The main clue to this will occur when some piece of data is perceived that does not fit with expectations based on the internal model. When a person's expectations do not match with what is perceived, this conflict can be resolved by (a) adopting a new model, (b) revising the existing model, or (c) changing one's goals and plans to accommodate the new situation (Manktelow and Jones, 1987). The inappropriate choice could easily sabotage SA efforts for quite some time.

If the new data can be incorporated into the model, it may merely indicate that a new situation (state of the model) is present which calls up different goals and plans accordingly. If the new data cannot be easily fit into the existing model, the model may be revised. A common problem will be to continue to revise the existing model to account for the new data when an alternate model is more appropriate. Something about the data must flag that a different situation is present which calls up the alternate model. Without this flag, the person may persist in a representational error whereby the data continues to be misinterpreted in light of the wrong model. Of course, if no appropriate new model exists, even when it is

recognized that the existing model is inadequate, there may still be significant errors while a new model is developed.

The relationship between SA and performance, although not always direct, can also be predicted. In general, it is expected that poor performance will occur (a) due to incomplete or inaccurate SA, (b) when the correct action for the diagnosed situation is not known or calculated, or (c) when time or some other factor limits a person's ability to carry out the correct action. For instance, Endsley (1990) found that SA was significantly related to performance only for those subjects who had the technical and operational capabilities to take advantage of such knowledge. The same study also found that poor SA would not necessarily lead to poor performance if subjects realized their lack of SA and were able to modify their behaviors to reduce the possibility of poor performance. Good SA can therefore be viewed as a factor which will increase the *probability* of good performance, but cannot necessarily guarantee it.

Future Directions for SA Research

In conclusion, an approach to decision making that takes situation awareness into account has been presented that can be of utility to researchers in a variety of arenas. The approach, emphasizing a descriptive view of expert human decision making in real world dynamic environments, is quite different from normative views of decision making developed from static laboratory based tasks. The concern shifts from the application of analytic rules (the use of which shifts from task to task) to a more holistic view which emphasizes the object of a decision maker's concern – classifying the situation. This shift in emphasis leads to the goal of designing systems that will (a) support SA requirements in complex environments, and (b) provide *triggers of the appropriate schema* to support decision making.

Design

A marked difference occurs in the shift in emphasis from providing human-machine interfaces (HMI) which present all needed *data*, to HMI which provide needed *SA*. This shift encompasses the fact that operators of complex systems need information that is organized according to goals, as opposed to disparate bits of data that must then be transformed and integrated. HMI that provide SA will simultaneously support multiple operator goals at a minimum workload, taking into consideration constraints of attention, the failings of human sampling and the debilitating effects of stressors. While some guidelines have been compiled for this (Endsley, 1988a), a considerable amount of both basic and applied research is needed to meet this goal in HMI system design.

Carefully constructed research paradigms are needed to ensure that researchers do not succeed in increasing SA on some aspects at the expense of other aspects. Furthermore, bearing human failings and constraints in mind, the development of HMIs which will support higher level SA needs (integration and projection) requires far more research. Artificial systems for status projection (often billed as decision support systems), for instance, need to be effectively integrated and evaluated to determine if they will provide any real benefit over

existing human capabilities. Current efforts at HMI development that automatically filter out all but system determined high priority information need to be examined carefully to ensure that other global SA needs required for rapid shifts between goals do not get lost. For the most part, the huge number of technologies being considered for HMI designs (both hardware and software) should be carefully evaluated to determine if they really benefit operator SA. Only by looking at the impact on SA needs as a whole can this determination be made, as a particular technology may improve SA in some areas, but only at the expense of SA in other areas.

Training

The goal of improving pilot situation awareness can be met by incorporating SA into training programs in several ways. (See Endsley (1989c) for a detailed discussion). First, SA-oriented training programs can be developed that instruct operators in the components of important schema, the dynamics and functioning of system components and projection of future actions based on these dynamics. This type of SA-oriented training is greatly needed in many areas to supplement traditional technology-oriented training.

Secondly, it should be recognized that SA is not a passive process. Operators must actively work to achieve it. The skills required for achieving and maintaining good SA need to be identified and formally taught in training programs.

Thirdly, feedback is an important component of the learning process. Feedback on the accuracy and completeness of operator SA could be incorporated into training programs to allow operators to understand their mistakes and better assess and interpret the environment, leading to the development of more effective sampling strategies and better schema for integrating information.

Construct Exploration

There is much that is still not known about situation awareness. At a minimum, exact SA requirements need to be determined for various systems. This has been accomplished for air-to-air fighter aircraft (Endsley, in press) and advanced bombers (Endsley, 1989a), but for many other types of systems, designers are working with only simple information requirements without an understanding of how the information needs to be integrated to support operator functions or of overall SA needs.

In addition, more basic questions remain. How is higher level SA generated from lower level concepts? What is the relative importance of each level? What are the really vital cues which trigger important schema? What are these mental models which are key components of the decision making process? Situation models, being a virtual reflection of system models, may serve to shed some light on this type of research. If mental models are truly "mechanisms whereby humans are able to generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future states" as described by Rouse and Morris (1985), then three of the four criteria (system functioning, states and predictions) can be determined by examining situation models across contexts. This type of effort may serve to help create a better understanding of the nature of mental models.

Finally, very little is really known about individual differences in SA. While there is much speculation that some individuals are clearly superior at obtaining and maintaining SA in complex environments, evidence for this is anecdotal and virtually nothing is known about the factors that may make one person better at SA than another. Recent research (Endsley and Bolstad, in preparation) indicates that such individual differences may be attributable to a combination of abilities in areas such as perception, memory, time sharing, spatial abilities and personality factors. More research along these lines is clearly needed to direct efforts in HMI design and training. By learning more about SA requirements and the SA construct as a whole, more effective HMI design and training programs can be established to support decision making in complex environments.

References

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*(1), 33-53.

Bacon, S. J. (1974). Arousal and the range of cue utilization. *Journal of Experimental Psychology, 102,* 81-87.

Baddeley, A. D. (1972). Selective attention and performance in dangerous environments. *British Journal of Psychology, 63,* 537-546.

Barber, P. J., & Folkard, S. (1972). Reaction time under stimulus uncertainty with response certainty. *Journal of Experimental Psychology, 93,* 138-142.

Bartlett, F. C. (1943). Fatigue following highly skilled work. *Proceedings of the Royal Society (B), 131,* 147-257.

Beach, L. R., & Mitchell, T. R. (1987). Image theory: Principles, goals and plans in decision making. *Acta Psychologica, 66,* 201-219.

Bettman, J. R., & Kakkar, P. (1977). Effects of information presentation format on consumer information acquisition strategies. *Journal of Consumer Research, 3,* 233-240.

Biederman, I., Mezzanotte, R. J., Rabinowitz, J. C., Francolin, C. M., & Plude, D. (1981). Detecting the unexpected in photo interpretation. *Human Factors, 23,* 153-163.

Broadbent, D. E. (1954). Some effects of noise on visual performance. *Quarterly Journal of Experimental Psychology, 6,* 1-5.

Broadbent, D. E. (1971). *Decision and stress.* London: Academic Press.

Broadbent, D. E., & Gregory, M. (1965). The effects of noise and signal rate upon vigilance analyzed by means of decision theory. *Human Factors, 7,* 155-162.

Callaway, E., III, & Dembo, D. (1958). Narrowed attention: A psychological phenomenon that accompanies a certain physiological change. *AMA Archives of Neurology and Psychiatry, 79,* 74-90.

Carmino, A., Idee, E., Larchier Boulanger, J., & Morlat, G. (1988). Representational errors: Why some may be termed diabolical. In L. P. Goodstein, H. B. Anderson, and S. E. Olsen (Eds.), *Tasks, errors and mental models* (pp. 240-250). London: Taylor and Francis.

Casson, R. W. (1983). Schema in cognitive anthropology. *Annual Review of Anthropology, 12,* 429-462.

Christensen-Szalanski, J. J. J. (1978). Problem solving strategies: A selection mechanism, some implications, and some data. *Organizational Behavior and Human Performance, 22,* 307-323.

Christensen-Szalanski, J. J. J. (1980). A further examination of the selection of problem-solving strategies: The effects of deadlines and analytic aptitudes. *Organizational Behavior and Human Performance, 25,* 107-122.

Damos, D., & Wickens, C. D. (1980). The acquisition and transfer of time-sharing skills. *Acta Psychologica, 6,* 569-577.

Davis, D. R. (1948). *Pilot error.* London: HMSO.

Davis, E. T., Kramer, P., & Graham, N. (1983). Uncertainty about spatial frequency, spatial position, or contrast of visual patterns. *Perception and Psychophysics, 5,* 341-346.

deGroot, A. (1965). *Thought and choice in chess.* The Hague, The Netherlands: Mouton.

Dreyfus, S. E. (1981). *Formal models vs. human situational understanding: Inherent limitations on the modeling of business expertise* (ORC 81-3). Berkeley: Operations Research Center, University of California.

Endsley, M. R. (1987a). *SAGAT: A methodology for the measurement of situation awareness* (NOR DOC 87-83). Hawthorne, CA: Northrop Corporation.

Endsley, M. (1987b). The application of human factors to the development of expert systems for advanced cockpits. *Proceedings of the Human Factors Society 31st Annual Meeting* (pp. 1388-1392). Santa Monica, CA: Human Factors Society.

Endsley, M. R. (1988a). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97-101). Santa Monica, CA: Human Factors Society.

Endsley, M. (1988b). Situation Awareness Global Assessment Technique (SAGAT). *Proceedings of the National Aerospace and Electronics Conference (NAECON)*. New York: IEEE.

Endsley, M. R. (1989a). *Final report: Situation awa eness in an advanced strategic mission* (NOR DOC 89-32). Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (1989b). A methodology for the objective measurement of situation awareness. In *Situational Awareness in Aerospace Operations* (pp. 1/1 - 1/9). Copenhagen, Denmark: NATO - AGARD.

Endsley, M. R. (1989c). Pilot situation awareness: The challenge for the training community. *Proceedings of the Interservice/Industry Training Systems Conference (I/ITSC)* (pp. 111-117). Ft Worth, TX: American Defense Preparedness Association.

Endsley, M. R. (1990). Predictive utility of an objective measure of situation awareness. *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 41-45). Santa Monica, CA: Human Factors Society.

Endsley, M. R. (in press). A survey of situation awareness requirements in air-to-air combat fighters. *International Journal of Aviation Psychology.*

Endsley, M. R., & Bolstad, C. A. (in preparation). *Individual differences in pilot situation awareness.* Manuscript in preparation.

Eysenck, M. W. (1982). *Attention and arousal: cognition and performance.* Berlin: Springer-Verlag.

Federal Aviation Administration. (1990). *The National Plan for Aviation Human Factors.* Washington: U.S. Department of Transportation.

Fracker, M. L. (1987). *Situation awareness: A decision model.* Unpublished manuscript. Dayton, OH.

Fracker, M. L. (1988). A theory of situation assessment: Implications for measuring situation awareness. *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 102-106). Santa Monica, Ca: Human Factors Society.

Fracker, M. L. (1989). Attention gradients in situation awareness. In *Situation Awareness in Aerospace Operations* (pp. 6/1-1/10). Copenhagen, Denmark: NATO - AGARD.

Hamm, R. M. (1988). Moment-by-moment variation in expert's analytic and intuitive cognitive activity. *IEEE Transactions on Systems, Man and Cybernetics, 18*(5), 757-776.

Hammond, K. R. (1986). *A theoretically based review of theory and research in judgment and decision making* (CRJP 260). Boulder, CO: Center for Research on Judgment and Policy, University of Colorado.

Hammond, K. R. (1988). *Judgment and decision making in dynamic tasks* (CRJP 282). Boulder, CO: Center for Research on Judgment and Policy, University of Colorado.

Herstein, J. A. (1981). Keeping the voter's limit in mind: A cognitive processing analysis of decision making in voting. *Journal of Personality and Social Psychology, 40,* 843-861.

Hinsley, D., Hayes, J. R., & Simon, H. A. (1977). From words to equations. In P. Carpenter and M. Just (Eds.), *Cognitive processes in comprehension* Hillsdale, N.J.: Erlbaum.

Hockey, G. J. R. (1970). Effect of loud noise on attentional selectivity. *Quarterly Journal of Experimental Psychology, 22,* 28-36.

Hockey, G. R. J. (1986). Changes in operator efficiency as a function of environmental stress, fatigue and circadian rhythms. In K. Boff, L. Kaufman, and J. Thomas (Eds.), *Handbook of perception and performance* (pp. 44/1- 44/49). New York: John Wiley.

Holland, J. H., Holyoak, K. F., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning and discovery.* Cambridge: MIT Press.

Humphreys, G. W. (1981). Flexibility of attention between stimulus dimensions. *Perception and Psychophysics, 30,* 291-302. *Intraflight Command, Control and Communications Symposium Final Report.* (1986). Nellis AFB, NV: 57th Fighter Wing.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306-340.

Janis, I. L. (1982). Decision making under stress. In L. Goldberger and S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects* (pp. 69-87). New York: The Free Press.

Jerison, H. J. (1957). Performance on a simple vigilance task in noise and quiet. *Journal of the Acoustical Society of America, 29,* 1163-1165.

Jerison, H. J. (1959). Effects of noise on human performance. *Journal of Applied Psychology, 43,* 96-101.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Keinan, G. (1987). Decision making under stress: Scanning of alternatives under controllable and uncontrollable threats. *Journal of Personality and Social Psychology, 52*(3), 639-644.

Keinan, G., & Friedland, N. (1987). Decision making under stress: Scanning of alternatives under physical threat. *Acta Psychologica, 64,* 219-228.

Kellog, R. T. (1980). Is conscious attention necessary for long-term storage? *Journal of Experimental Psychology: Human Learning and Memory, 6,* 379-390.

Klein, G. A. (1989a). Do decision biases explain too much? *Human Factors Society Bulletin, 32*(5), 1-3.

Klein, G. A. (1989b). Recognition-primed decisions. In W. B. Rouse (Eds.), *Advances in man-machine systems research* (pp. 47-92). Greenwich, Conn: JAI Press, Inc.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground. *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 576-580). Santa Monica, CA: Human Factors Society.

Kuhn, T. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

LaBerge, D., VanGelder, P., & Yellot, S. (1971). A cueing technique in choice reaction time. *Journal of Experimental Psychology, 75,* 321-328.

Lipshitz, R. (1987). *Decision making in the real world: Developing descriptions and prescriptions from decision maker's retrospective accounts.* Boston, MA: Boston University Center for Applied Sciences.

Logan, G. D. (1988). Automaticity, resources and memory: Theoretical controversies and practical implications. *Human Factors, 30*(5), 583-598.

Mandler, G. (1979). Thought processes, consciousness and stress. In V. Hamilton and D. M. Warburton (Eds.), *Human stress and cognition: An information-processing approach* Chichester: Wiley and Sons.

Mandler, G. (1982). Stress and thought processes. In L. Goldberger and S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects* (pp. 88-104). New York: The Free Press.

Manktelow, K., & Jones, J. (1987). Principles from the psychology of thinking and mental models. In M. M. Gardiner & B. Christie (Eds.), *Applying cognitive psychology to user-interface design* (pp. 83-117). Chichester: Wiley and Sons.

Martin, M., & Jones, G. V. (1984). Cognitive failures in everyday life. In J. E. Harris and P. E. Morris (Eds.), *Everyday memory, actions and absent-mindedness* (pp. 173-190). London: Academic Press.

Mayer, R. E. (1983). *Thinking, problem solving, cognition.* New York: W. H. Freeman and Co.

Mintzburg, H. (1973). *The nature of managerial work.* New York: Harper and Row.

Mosier, K. L., & Chidester, T. R. (1991). Situation assessment and situation awareness in a team setting. In Y. Queinnec & F. Daniellou (Eds.), *Designing for Everyone* (pp. 798-800). London: Taylor and Francis.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century, Crofts.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231-259.

Noble, D., Boehm-Davis, D., & Grosz, C. (1987). *Rules, schema and decision making.* (NR 649-005). Engineering Research Associates.

Norman, D. A. (1976). *Memory and attention.* New York: Wiley and Sons.

Norman, D. A. (1983). A psychologist reviews human processing: Human errors and other phenomena suggest processing mechanisms. In *Five Papers on Human-Machine Interaction* La Jolla, CA: University of California, San Diego.

National Transportation Safety Board. (1973). *Aircraft Accident Report: Eastern Airlines 401/L-1011, Miami, Florida, December 29, 1972.* Washington, D.C.: Author.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory and Cognition, 3,* 519-526.

Posner, M. I., Nissen, J. M., & Klein, R. (1976). Visual dominance: An information processing account of its origin and significance. *Psychological Review, 83,* 157-171.

Posner, M. I., Nissen, J. M., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. L. Pick and E. J. Saltzman (Eds.), *Modes of Perceiving and Processing* (pp. 137-157). Hillsdale, NJ: Erlbaum Associates.

Press, M. (1986). *Situation awareness: Let's get serious about the clue-bird.* Unpublished manuscript.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. New York: North-Holland.

Reason, J. (1984). Absent-mindedness and cognitive control. In J. E. Harris and P. E. Morris (Eds.), *Everyday memory, action and absent-mindedness* (pp. 111-132). London: Academic Press.

Reason, J. (1987). A framework for classifying errors. In J. Rasmussen, K. Duncan, and J. Leplat (Eds.), *New technologies and human error* (pp. 5-14). Chichester: Wiley and Sons.

Reason, J. (1988). Framework models for human performance and error: A consumer's guide. In L. P. Goodstein, H. B. Andersen, and S. E. Olsen (Eds.), *Tasks, errors and mental models* (pp. 35-49). London: Taylor and Francis.

Roschelle, J., & Greeno, J. G. (1987). *Mental models in expert physics reasoning* (DTIC #AD-A184 106). Berkeley, CA: University of California.

Rouse, W. B., & Morris, N. M. (1985). *On looking into the black box: Prospects and limits in the search for mental models* (DTIC #AD-A159080). Atlanta, GA: Center for Man-Machine Systems Research, Georgia Institute of Technology.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding.* Hillsdale, NJ: Erlbaum.

Sharit, J., & Salvendy, G. (1982). Occupational stress: Review and reappraisal. *Human Factors, 24*(2), 129-162.

Sheridan, T. (1981). Understanding human error and aiding human diagnostic behavior in nuclear power plants. In J. Rasmussen and W. B. Rouse (Eds.), *Human detection and diagnosis of system failures* New York: Plenum Press.

Shugan, S. M. (1980). The cost of thinking. *Journal of Consumer Research, 7,* 99-111.

Sieber, J. E. (1974). Effects of decision importance on ability to generate warranted subjective uncertainty. *Journal of Personality and Social Psychology, 30*(5), 688-694.

Sundstrom, G. A. (1987). Information search and decision making: The effects of information displays. *Acta Psychologica, 65,* 165-179.

Sweller, J. (1988). Cognitive load during problem solving: Effects on Learning. *Cognitive Science, 12,* 257-285.

Treisman, A., & Paterson, R. (1984). Emergent features, attention and object perception. *Journal of Experimental Psychology: Human Perception and Performance, 10*(1), 12-31.

Tulving, E. (1972). Episodic and semantic memory. In Tulving and Donalson (Eds.), *Organization of memory* (pp. 385-387). New York: Academic Press.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40,* 385-398.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(30 January), 453-458.

VanDijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Wachtel, P. L. (1967). Conceptions of broad and narrow attention. *Psychological Bulletin, 68,* 417-429.

Weltman, G., Smith, J. E., & Egstrom, G. H. (1971). Perceptual narrowing during simulated pressure-chamber exposure. *Human Factors, 13,* 99-107.

Wickens, C. D., Stokes, A. F., Barnett, B., & Hyman, F. (1988). Stress and pilot judgment: An empirical study using MIDIS, a microcomputer-based simulation. *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 173-177). Santa Monica, CA: Human Factors Society.

Wickens, C. S. (1984). *Engineering psychology and human performance.* Columbus, Ohio: Charles E. Merrill Publishing Co.

Wirstad, J. (1988). On knowledge structures for process operators. In L. P. Goodstein, H. B. Anderson, & S. E. Olsen (Eds.), *Tasks, errors and mental models* (pp. 50-69). London: Taylor and Francis.

Woodhead, M. M. (1964). The effects of bursts of noise on an arithmetic task. *American Journal of Psychology, 77,* 627-633.

Wright, P. (1974). The harassed decision maker: Time pressures, distractions, and the use of evidence. *Journal of Applied Psychology, 59* (5), 555-561.

Situation Awareness is Adaptive, Externally-Directed Consciousness

Kip Smith & P.A. Hancock

University of Minnesota

We define situation awareness (SA) as adaptive, externally-directed consciousness. This definition dispels an artificial and contentious claim evident in the literature, namely that SA is either exclusively knowledge or exclusively process. This misguided rivalry has more to do with alternative stances toward the study of human behavior rather than with SA *per se*. By defining SA as an aspect of consciousness, we are able to clarify two key issues: criteria for evaluating performance at SA and the locus of competence for SA. We find the source of goals and performance criteria for SA to be a normative arbiter in the task environment. We ascertain competence at SA to be the invariant at the core of an adapted agent's Neisser cycle (Neisser, 1976). We introduce an observer construct, the 'risk-space,' that embodies competence at SA. The risk-space generates up-to-the-minute knowledge and drives actions that satisfy the goals and performance criteria specified in the task environment.

Introduction

Like stress and attention, SA is something we all recognize. However, we have yet to settle on exactly what we collectively think SA is. This is not the agreement of politicians to disagree but rather a tacit recognition that our understanding is still incomplete. Definitions of SA have emphasized either knowledge (e.g., Hopkin, this volume) or process (e.g., Endsley, 1988; this volume). This disparity of definition highlights the duality that present theory bequeaths to SA. As knowledge, SA is up-to-the-minute comprehension of task-relevant information that enables appropriate decision making under stress. As cognition-in-action (Lave, 1988), SA fashions behavior in anticipation of the task-specific consequences of alternative actions. We contend that this apparent contradiction is dispelled by understanding SA in terms of the Neisser cycle (Neisser, 1976). We extend the account of Tenney, Adams, Pew, Huggins, and Rogers (1992) to propose that: (i) SA is best defined as adaptive, externally-directed consciousness; and (ii) SA is the invariant at the core of an adapted agent's Neisser cycle that generates both up-to-the-minute knowledge and action that anticipates signals in the task environment.

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

Situation Awareness Defined

In defining situation awareness as *adaptive, externally-directed consciousness*, we take consciousness to be that part of an agent's knowledge-generating behavior that is within the scope of intentional manipulation. As shown in Figure 1, we view SA as purposeful behavior that is directed toward achieving a goal in a specific task environment. It has as its products (i) knowledge about, and (ii) directed action within that environment. We argue that SA is more than performance (knowledge and behavior) in the task environment. More fundamentally, it is the invariant capacity to direct consciousness to generate competent performance given any particular situation that might unfold. In what follows, we present a hypothesis about the structure of the knowledge that enables SA in air-traffic controllers.

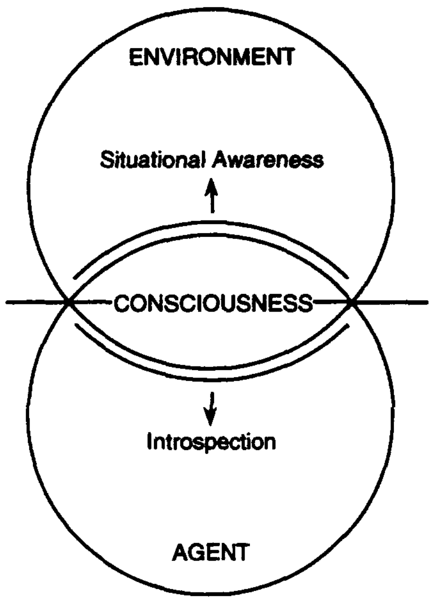


Figure 1. An approach to defining situation awareness (SA) through explicit recognition of the centrality of externally oriented consciousness. The central [horizontal] line provides an arbitrary distinction between exogenous and endogenous orientations of consciousness and represents a distinction between SA and introspection. Details of the organization are given in the text.

We see a direct relationship between consciousness, SA, and adaptation. Adaptation is the match of the agent's knowledge, beliefs, and goals to the information and activity made available by the environment (Simon, 1981). In our definition, we intend the phrase "externally-directed" to indicate that SA is goal-driven behavior. The phrase reveals explicitly that we place the focus of this behavior in the environment rather than in the agent's

'head'. As goal-driven behavior, SA is amenable to the methodology of information-processing psychology (Wickens, 1993).[1]

SA, like adaptation, is a dynamic concept that exists at the interface between the agent and its environment and that requires an assessment of the environment-agent relationship to be understood. As adaptation, SA presumes extensive experience in an environment and the development of an armory of appropriate, alternative courses of action (Holland, 1992/1975). We submit that SA presumes adaptation of a particular kind. As expressed in aviation and as all skilled practitioners know, SA is all about having *the right stuff.* This notion implies complete and 'natural' adherence to task goals and to criteria for performance. This, in turn, implies the existence of a specification of the task the agent is to perform and of measures for evaluating that performance. To possess SA, to have the right stuff, the agent must necessarily have developed a level of adaptive capability sufficient to match the specification of task goals and of criteria for assessing performance variables. Thus, SA is adaptation to a singular source of constraint: a normative arbiter that defines the stuff that it right.

As shown in Figure 2, we see the arbiter and its dicta residing in the task environment. The real problem in the current formulations of SA is the failure to articulate the presence in the environment of normative specifications and criteria for the performance of the agent's task. While individuals may exhibit situated, outwardly-directed consciousness, it is not until the externally-defined task is made explicit that their behavior achieves the status we wish to reserve for SA. To qualify as SA, the agent first must intend its goals, beliefs, and knowledge to match the task and performance specified by dicta from its environment and, then, must succeed to some degree in meeting those expectations.

Failure to recognize the role of the normative arbiter of performance has clearly been a source of confusion in the literature on SA. Until an external task and criteria for its performance are specified, examination of greater or lesser degrees of SA or even of loss of SA remains problematic. If the agent were to dictate the goal, SA would always be perfect since whatever was perceived would be the goal. To borrow a phrase, we might call this 'ambient SA'. However, in human factors and indeed in most realms of human activity there are goals, either set by others, or set by ourselves at some previous point in time. Such goals are supported by task relevant cues and negated by task irrelevant cues. Once we accept the external arbiter as the key constraint on the adaptation of the agent, the directed nature of SA becomes a more manageable construct. Only with a specified task and concrete performance criteria can we begin to talk about how well adapted a particular agent is with respect to that environment.

As an emergent property of adaptation, SA is appropriately discussed within the framework of the ecological movement (e.g., Gibson, 1969). The ecological approach affirms the importance of the environment in dictating what goes on with behavior. The focus is on the agent's action as shaped by its interaction with its environment. Yet, despite its name, the ecological approach is more than a swing toward environmental determinism. The critical facet of understanding, predicated on Gibson's work, is a recognition that it is the *interaction* of the organism and its environment that provides meaningful insight into the actions an agent takes. (This is of course a highly simplified account of only a single facet of the ecological position.)

1 We do not deny that there is internally directed consciousness (e.g., introspection), but maintain that consciousness directed to internal representations (e.g., mental models) is a meta-construct that leads to a number of philosophical polemics that fail to help resolve current practical questions about SA.

As it is the relationship of organism and environment that is critical, statements about the situation *alone*, or about awareness *alone*, are liable to be much less substantive than those about emergent properties derived from an interactive view. Given this perspective, any attempt at comprehension of SA without a viable understanding of situations would be, to say the least, difficult. The submission of the ecologists would be that if we can continue to study SA with the techniques that focus overwhelmingly on the agent as the individual 'unit of concern', the critical emergent properties are as likely to 'emerge' as the Encyclopedia Brittannica from the combined efforts of Eddington's anthropoid typists.

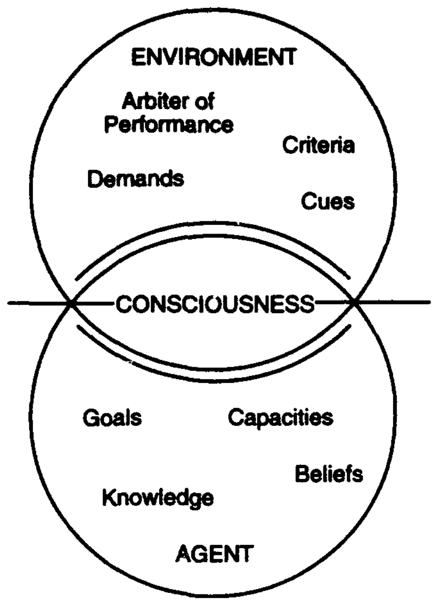


Figure 2: Constraints on SA. The singular constraint is the presence of a normative arbiter of performance in the agent's task environment. The arbiter specifies for the agent task-relevant goals and criteria for performance. Adaptation to the environment requires the agent to adopt the arbiter's specification of goals and performance variables. Cues and demands are stimuli that unfold in the environment. The agent's internal constraints are those that shape its intentionality.

To return to our definition, an agent's SA is behavior developed to generate knowledge and action given the structure of its environment and the goals and performance norms specified in that environment. The explicit use of the conditional term 'given' reveals our conviction that the root of SA is adaptive behavior that changes in response to alternative situations. [illegible]mportant to note that this definition denies a claim of SA to any agent that is merely c[illegible] and attending to its environment. Further, it denies a claim of SA to an agent that [illegible]e fully capable of SA but that is not actively the goal specified by the arbiter of performance. Rather, to stake a claim to SA, an agent must be actively seeking information and taking action; its consciousness must be outward-directed and constrained by goals established by the situation that informs its search. Without the constraint of

externally-defined goals, without an external arbiter of behavior, SA degenerates to mere passive observation.

By defining SA as a generative process of knowledge creation and informed action-taking, we expressly deny that SA is merely a snap-shot of a 'mental model'. Rather, SA is the process of its constitution. The experience of air-traffic controllers 'losing the picture' illustrates this point. As Hopkin (this volume) recounts, the controllers' job is to construct knowledge of their sector and to take action on the basis of this knowledge. They call the knowledge they generate 'the (big) picture'. At times, they 'lose the picture' – their knowledge becomes insufficient to support their task. Experienced controllers on the job are often sufficiently self-aware to recognize they are losing the picture as it happens. This 'meta-knowledge' argues our point: it is their SA that builds 'the picture' and that enables them to know that what they know is not up to the task they face. SA not only supports the construction of 'the picture' but also assesses its integrity.

Performance, Competence, and SA

The distinction between competence and performance is a persistent theme in studies of cognition (Anderson, 1990; Chomsky, 1965; Marr, 1982). Performance is action situated in the world. Competence is knowledge that supports behavior but is independent of the situation. Performance is contingent upon information made available by the environment; competence is invariant of the particulars of a situation. The utility of this distinction is the leverage it provides to understanding the performance that we observe – the normatively focused knowledge-generation and action-taking that characterize SA.

An analysis of competence asks a simple question: What is the problem that this agent's behavior is the solution for? Specification of the problem focuses on the agent's knowledge, its goals, the information available in the environment, and the actions the agent may take to meet its goals. An analysis of competence is unconcerned with the actual processes (e.g., representations, mental models) that produce the agent's performance. As Newell (1982) pointed out, the resulting description places constraints on behavior, it does not prescribe performance. Full prescription or emulation of the agent's behavior requires a complete accounting of the agent's representation and process.

As behavior, SA is the solution to a problem. To paraphrase a member of the military aircraft industry, the problem is 'knowing what ya gotta know in order to do what ya gotta do'. The solution for the agent is to pay attention to those cues and demands in the environment that enable it to take action that aligns with the dicta of the arbiter of performance. Tenney and others (1992) propose that Neisser's (1976) perceptual cycle provides a framework for understanding how SA works, that is, how an agent and its environment interact in a manner to satisfy the arbiter by generating skilled performance. We agree. The Neisser cycle is reproduced in Figure 3. Information and action flow continuously around the cycle. Starting arbitrarily at the top, the environment informs the agent, modifying its knowledge. Knowledge directs the agent's activity in the environment. That activity samples and, perhaps, anticipates or alters the environment which, in turn, informs the agent. The informed, directed sampling and/or anticipation capture the essence of the performance of SA.

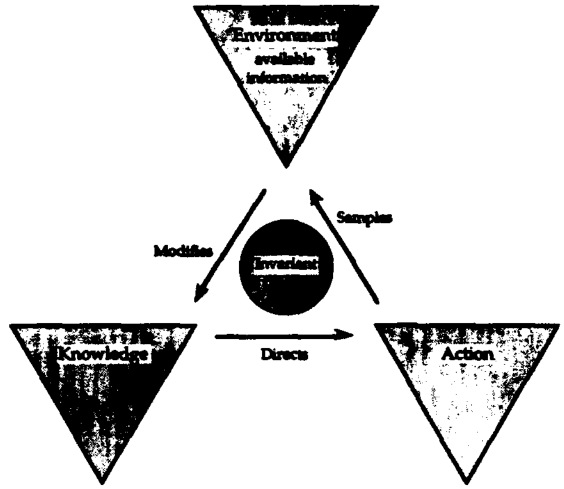


Figure 3. Neisser's (1976) perceptual cycle. The invariant at the core of the cycle specifies the agent's adaptation to it environment. It structures the information made available by the environment, the agent's knowledge and the actions the agent takes to meet the goals specified by the arbiter of performance.

To go beyond performance, to capture competence at SA, we include in Figure 3 the invariant that links the three elements of the Neisser cycle. The invariant is the structure of the agent's adaptation to the environment: it forms the linkage among information, knowledge, and action that produces competent behavior. Specifically, the invariant codifies the information that the environment may make available, the knowledge the agent requires to assess that information, and the action the knowledge will direct the agent to take to meet its goals. Our candidate representation for the invariant in air traffic control is a multidimensional 'risk-space' (Smith & Hancock, 1992). Figure 4 schematically presents two possible dimensions of the multidimensional ATC risk space. The risk space is a generalization of a mathematical formulation of human performance in monitoring tasks (Phatak & Bekey, 1969; Moray, 1986). The axes of the risk space for ATC are defined by factors in the environment that compromise safety. These factors define sources of information that the agent attends in order to satisfy the arbiter's norms for performance. Thresholds of safety parse the risk space into "decision regions". The thresholds are performance criteria defined by either the arbiter or the agent that differentiate alternative control decisions. Each decision region is associated with one, and only one, course of action.

Many factors define the ATC risk space. Two of them – aircraft separation and relative velocity – define the portion of the ATC risk space sketched in Figure 4. As shown in Figure 4, at every instant of time the separation and relative velocity between a pair of aircraft define a unique point in the risk space framework. The action to be taken by the controller is specified by the decision region which the aircraft occupies in the risk space. As illustrated in Figure 5, this point traverses the decision regions as aircraft fly through the airspace. The path that aircraft trace through the risk space can be used to anticipate future events but

remains independent of their physical locations in the airspace. As a result, the risk space facilitates prognosis but is itself invariant across airspace (sectors) and time.

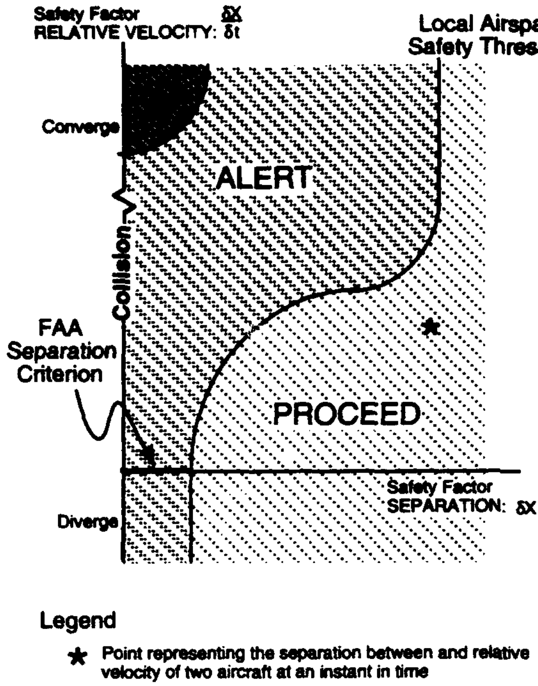


Figure 4. A simplified portion of the risk space framework for Air Traffic Control. The risk space is an observer construct that specifies the controller's adaptation and that can be used to describe and predict SA performance. The axis are defined by factors in the environment that compromise safety. Thresholds of safety parse the risk space into "decision regions" associated that specify the action to be taken by the controller.

The decision regions shown in Figures 4 and 5 reflect the risk space for today's environment where ATC issue commands. The region in the lower right labeled "proceed" is the region where separation is great and relative velocity is low: there is little danger of collision as long as ATC directives are followed. The goal of all control decisions is to keep all aircraft in the decision region labeled "proceed". Conversely, the small region in the upper left corresponds to situations where, we hypothesize, collision is imminent and pilots must take evasive action. The large region in the middle labeled "alert" is admittedly vastly over-simplified. It is the region where ATC issues control decisions that are intended to modify the distribution of aircraft in the sector.

The risk space framework integrates information critical to safety considerations in a manner that specifies the action appropriate for a given sector at a particular time. The decision regions specify (and can be used to communicate) the actions that pilots and controllers need to take given the unfolding of events in the airspace. The risk space is a observer construct: we do not claim that the adaptation of any individual controller

necessarily takes this form. Our claim is that the risk space framework (when fully developed) specifies ATC competence and, therefore, can be used to describe and predict the competent controller's SA performance.

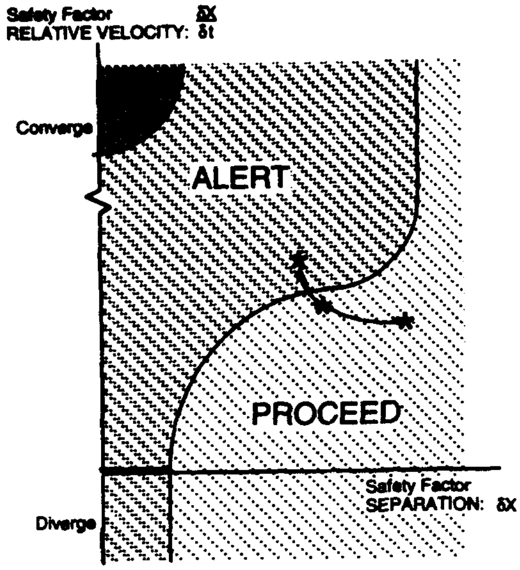


Figure 5. Motion through the risk space reflects flight through the airspace and indicates the appropriate action.

Discussion

We have argued that SA is a facet of consciousness. It is one in a long line of *energetic* constructs (e.g., attention and workload) which in a progressive fashion have re-introduced consciousness into scientific discussion of human behavior[2]. As consciousness has re-emerged, age-old questions as to its nature – product or process, knowledge or performance – have percolated to the surface. We submit that philosophical resolution of these questions is

2 This gradual rehabilitation of consciousness is in sequential response to Watson's (1913) [understandable] excision of the mental from what is a mentalistic science. Parenthetically, we see SA in the light of the argument between the tradition of information processing psychology and the nascent ecological approach to human factors. While we believe the ecological approach will eventually provide the predominant perspective for human factors, we paradoxically do not see the ecological approach as an eventual science of individual behavior. We suggest that their respective 'units of analysis' - the human for information processing psychology and the emergent elements of the human-machine-environment triad for the ecological approach - will eventually favor an approach in which human remain 'the hero of the story'. So, while 'blaming the victim' is inappropriate in human factors, 'cheering the hero' is mandatory in exploring human behavior. We believe that even scientists have not and probably cannot free themselves of this latter perspective.

unlikely be had in any facile manner. However, for *practical* purposes, we in human factors can impose some operational bounds that allow us to define our terms.

In this piece we have articulated such bounds. We posit first that SA is externally-directed towards a task environment. This means SA is facet of consciousness but not necessarily all of consciousness. Second, for SA we insist that goals and criteria for performance must be made explicit in the environment. SA is referenced to those goals and demands. Finally, we recognize SA as an invariant but adaptive component in a cycle of knowledge, action, and information. In this cycle, knowledge directs adaptive behavior that modifies the environment that then informs knowledge and so the cycle continues. Adaptive behavior that satisfies the arbiter of performance is a byproduct of a cycle that is driven by competence at SA.

We cannot apologize if our approach proves disturbing to those who wish static definitions. Any useful notion of SA must clearly face the central problems of generalization and non-stationarity posed by environments that are continually changing and by agents that are continually adapting and learning. The risk-space is an observer construct that, we contend, captures the structure of a skilled agent's adaptation to a dynamic environment filled with moving targets. The risk-space is defined by sources of information, by criteria for action, and by the actions themselves that together address the goals and standards for performance specified by a normative arbiter. The sources of information, criteria, and actions are invariant across situations; the risk-space is the structure that makes them operative. Like a grammar for language, the risk-space generates all possible situations and prescribes adaptive, appropriate behavior. The risk-space is a specification of competence at SA. It supports the cycle of adaptive behavior that we all recognize as SA.

In sum, we see value in the construct we are calling situation awareness. We expect to see it and other, even clearer manifestations of consciousness as we design and evaluate purposive human-machine-environment systems. That SA will prove a valuable step along this path is an important rationale in and of itself. Despite polemics about comparable and overlapping psychological constructs, we believe the present movement is in the right direction.

References

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Erlbaum.
Chomsky, N. (1965). *Aspects of a theory of syntax.* Cambridge, MA: MIT Press.
Endsley, M.R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society ,32.*
Endsley, M.R. (this volume).
Gibson, J.J. (1969). *The ecological approach to visual perception.* Boston, MA: Houghton-Mifflin.
Holland, J.H. (1992/1975). *Adaptation in natural and artificial systems.* Ann Arbor, MI: University of Michigan Press.
Hopkin, V.D. (this volume). Situational awareness in air traffic control.
Lave, J. (1988) *Cognition in practice.* Cambridge, UK: Cambridge University Press.
Marr, D. (1982) *Vision.* San Francisco, CA: W.H. Freeman.

Moray, N. (1986). Monitoring behavior and supervisory control. In K.R. Boff, L. Kaufman, and J.P. Thomas (Eds.) *Handbook of perception and human performance, 2, 40.* New York, NY: Wiley.

Neisser, U. (1976) *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco, CA: W.H. Freeman.

Newell, A. (1982) The knowledge level. *Artificial Intelligence, 18*, 87-127.

Phatak, A.V., & Bekey, G.A. (1969). Decision processes in the adaptive behavior of human controllers. *IEEE Transactions of Systems, Science and Cybernetics, 5*, 339-352.

Simon, H.A. (1981) *Sciences of the artificial. Second Edition*. Cambridge, MA: MIT Press.

Smith, K., & Hancock, P.A. (1992). Managing risk under time stress. *Proceedings of the Human Factors Society ,36*, 1019-1023.

Tenney, Y.J., Adams, M.J., Pew, R.W., Huggins, A.W.F., & Rogers, W.H. (1992) A principled approach to the measurement of situation awareness in commercial aviation. NASA Contractor Report 4451.

Watson, J.B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*, 158-167.

Wickens, C.D. (1993). Workload and situation awareness: An analogy of history and implications. *Insight,: The Visual Performance Technical Group Newsletter, 14, 4.*

Situation in Mind: Theory, Application and Measurement of Situational Awareness

R. M. Taylor & S. J. Selcon

RAF Institute of Aviation Medicine

Summary

We propose to discuss the issues within a "question and answer" framework. The following four principle questions are addressed:

1. What is the problem?
2. What is the relevant theory?
3. What are the consequences of the theory?
4. Do the theoretical predictions fit experience and the experimental data?

The aim is to give short answers with sufficient information for follow-up, as required.

The Problem?

Aircrew operators of advanced systems have complained of a lack of "situational awareness", e.g. USAF F15 and F16 aircrew. Poor situational awareness (SA) seems to be associated with accidents and incidents, and with reduced mission effectiveness. It is generally thought that "experts" have fewer problems with SA than "novices". Without defining expertise, this suggests that poor SA is probably at least partly a training issue. However, increasing difficulties with operator SA seem to be associated with employing advanced automation and display/control technology in increasingly complex systems and highly dynamic environments. There is uncertainty about how designers should employ advanced technology to automate tasks and reduce unwanted operator workload, whilst at the same time providing the operator with the SA necessary to perform the intended operator functions and tasks. Also, there is uncertainty about how best to present the necessary situational information indirectly on synthetic and symbolic displays, with the trend towards more remote and less directly observable air warfare. The problem of SA arises from the continuing need for human involvement in advanced, highly automated aircraft systems.

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

Maintaining SA in advanced military aircraft is likely to be a problem simply because of the highly complex and dynamic nature of aircrew tasks. However, problems also seem to have arisen when the performance of previously "manual" tasks has been transferred to automation technology without due consideration of the impact on the operator's remaining functions and tasks. It seems that the operator can be taken outside of the information processing and decision-making loop, from where it can be difficult, if not impossible, to follow and monitor what is going on. This can present serious difficulties for operators of new systems, when there is a requirement to develop operator trust in the correct functioning of the automation. Also, it can present difficulties when the operator is required to intervene and take control in unexpected and abnormal situations and when required to initiate remedial action in emergencies, following say, enemy action, automation failure, or breakdown. Operator errors derived from erroneous assessments of situations are a price of automation.

The first problem for most human factors researchers interested in SA, many of whom will have heard SA described as "the buzz-word of the 1990's", is that of definition; i.e. what does it mean? Various definitions have been proposed by researchers. Concern has been expressed about the lack of a commonly accepted definition for what is regarded as a critical concept (Fracker 1988; Sarter & Woods 1991). This concern about definition may arise because existing definitions of SA are either too general to satisfy specific needs, or too specific to be generalizable to different domains. A central consideration is that SA is a state, i.e. a product rather than a process. Situation assessment is the appropriate term for the process that leads to SA. For this reason, definitions of SA are more likely to identify the requisite task knowledge, rather than to describe the mechanisms of knowledge acquisition and cognition associated with the situation assessment. Most definitions make reference to the operator's knowledge or understanding of the situation. Accessibility, comprehensiveness, cohesiveness and currency are attributes of cognitive representations that are claimed to be associated with SA. Distinctions are often made between the kinds of knowledge that are relevant to the operator's tasks. These include knowledge about the internal functioning of the system (e.g. the aircraft's avionics), understanding of the external environment in which the system is operating, and of the relationship between internal and external variables. Anticipation and thinking ahead seem to be particularly important in highly dynamic situations. It also seems important to know what is not known, or to be confident of how much uncertainty there is about. These are apparently contradictory, meta-knowledge (knowledge about the use of knowledge) requirements of good SA.

In research at the RAF IAM, the problem of definition has been circumvented by asking aircrew what factors were associated with examples, provided by aircrew, of good and poor SA (Taylor 1989). The relevant factors or constructs were identified by using the repertory grid technique, which is an interview technique developed by psychologists for eliciting knowledge of personal constructs without influencing that knowledge. The constructs associated with situational awareness, obtained from RAF aircrew in this way, concerned factors affecting: (a) the demand for attentional resources arising from the situation (complexity, variability, instability); (b) the supply of attentional resources in response to the situation demands, (arousal, spare mental capacity, concentration, division of attention); and (c) the resultant understanding or knowledge of the situation (familiarity, information quantity, information quality). Thus, for RAF aircrew at least, SA is thought to be a product of situational, attentional, and cognitive factors. A comprehensive definition of SA should refer to all three areas.

The Theory?

Many existing psychological theories are associated with SA, and can be considered to be relevant. But there is no integrated, global theory of SA. Theory of "awareness" tends to be treated separately from the theory of "situations". What is needed is a general theory that deals with what matters in situations and addresses how that knowledge is communicated, understood and acted upon. Physical theory must be relevant, since situational awareness concerns knowledge of, and interaction with, the real world. Theory of military "situations" is particularly relevant to the problems of the F15 and F16 fighter pilot. Presumably, the relevant theory is embodied in the military and political doctrine of air warfare, in the modelling of scenarios and missions, and in tactical philosophy and rules of engagement etc.

Most psychological theories of attention and cognition are potentially relevant, particularly in their application to human-machine systems interaction. But there is uncertainty over whether existing theories are sufficient, or whether additional theory, or a more integrated theory, is needed to resolve the problem of enhancing SA. It is outside the scope of this paper to discuss all potentially relevant theory. The focus here will be on the adequacy of psychological theory. Psychological theories that have been most readily applied to understanding SA are the following:

a. Theory of limited attentional resources
b. Theory of short term, working memory
c. Theory of long term memory, in particular Schemata Theory of knowledge structures

Discussions of the application of these theories to SA have been provided by Taylor 1987, Fracker 1988, Endsley 1988, and others. Additional potentially relevant psychological theories that could be useful to discussions of SA include the following:

- Theory of attentional priority, such as the race model theory of visual attention, which proposes mechanisms for visual recognition and attentional selection based on filtering, pigeonholing and queuing of high priority items (Yantis, & Johnson 1990; Bundesen 1990).

- Theory of perceptual organisation and structuring, such as feature integration theory of pre-attentive vision (Triesman & Gelade 1980), and theory of local and global levels of processing (Kinchla, 1980) which proposes mechanisms or rules for reducing the processing load in perception.

- Perceptual control theory of behaviour, such as the theory of layered protocols, which develops, in the context of human-machine interaction, the notion that all behaviour is directed to the control of perceptions at a variety of levels of abstraction (Taylor 1988).

- Theory of mental models which concerns how information is structured into a useable internal representation to solve problems (Johnston-Laird 1983).

- Theory of semantic memory, used originally to account for prose comprehension, which can be used to quantify the cognitive quality of displays in terms of semantic networks with nodes, links and gates (Chechile, Egglestone, Fleishman, & Sasseville, 1989).

- Theory of spatial orientation, spatial vision, visualisation, and visuo-spatial ability, which concerns individual differences and skills which may be associated with SA (Howard 1982).

- Theory of dynamic mental representations which concerns the acquisition and representation of dynamic information (Freyd 1987).

- Theory of human error which describes human performance in realistic tasks in terms of skill, rule, and knowledge-based behaviour. (Rasmussen 1986).

- Theory of naturalistic decision making in which situation assessment under time pressure is treated as a pattern recognition process, with mental simulation to test the consequences of proposed actions (Klein, Orsanu, & Calderwood, 1992).

- Theory of plans, goals and "situated acts", which deals with the differences between planned and opportunistic behaviour (Schank & Abelson 1977).

The Consequences?

SA is a broad concept. It draws upon a large area of psychological theory, and depends on interaction with the real world. Current theory is almost entirely psychological. It involves a limited selection from the range of psychological theory available, which may or may not be the most appropriate. Also, there is uncertainty about how adequately the theory deals with the interaction with the real world. An integrated theory would need to encompass both psychological and environmental variables.

Classification into the "who, what, when and where" of situations is part of the requirement. But it may not be sufficient to identify what knowledge is actually activated in "good" situational awareness. Meta-knowledge about the uncertainties, interdependencies, interactions and dynamics of situational variables is important in anticipating problems and thinking ahead. Understanding of the demands of the environment at this level of complexity is not well-integrated with psychological theory of "awareness", at least sufficiently to be able to make useful, testable predictions. Theory does not yet enable us to predict successfully what will be the awareness of an individual in a given situation. Arguably, experienced instructors and trainers can make valid predictions about individuals and situations they are familiar with, and presumably do so regularly in certifying people to do difficult jobs like flying. However, the frequent failure of warning systems to draw attention to and to lead to the solution of urgent problems in aircraft emergencies provides testimony to the weakness of our understanding of SA, as far as system engineering and design are concerned.

The primary function of theory is to make predictions and to produce testable hypotheses. Findings from tests of predictions and hypotheses should be used to accept or reject

hypotheses, and the lessons learnt should be used to refine the theory. The value of a theory of SA will need to be judged by the quality of the testable hypotheses arising from that theory. Since the theory of SA appears selective and fragmented, the consequences are that, as yet, there are few generalizable conclusions or predictions that seem to be more than common sense at best, and self-fulfilling prophesies at worst. One example is the circularity in the reasoning that, since SA involves memory processes, it can be measured by the accuracy of recall, and that good recall will be indicative of good SA. Related predictions derive from the notion that SA is associated with knowledge. These are that since knowledge can be improved by experience and training, the same should be so for SA and that differences between novices and experts should be indicative of improving SA. A prediction from the limited resource model of attention is that if the supply of attentional resources fails to match the demands in the situation, a breakdown of SA will occur, both in conditions of underload (boredom?) and overload. Knowledge also might be considered usefully as a limited resource, with associated supply and demand issues. However, it is difficult to conceive of examples of where too much knowledge is a disadvantage for SA, except perhaps in an instructional situation where a relatively high level of expert knowledge may make it difficult to communicate and to teach effectively.

More novel predictions might arise from the large number of potentially relevant theories in cognitive psychology. In particular, it might be more helpful to have a better understanding of the relationship between knowledge structures and the mechanisms of attentional priority, since these mechanisms presumably govern attentional control, allocation strategy, and management. Traditional perceptual theory treats perceptions as dynamic hypotheses. Perceptual control theory (PCT) of behaviour, advocated by Martin Taylor at DCIEM, deals explicitly with interaction between the observer and the real world (Taylor 1988). Under PCT, all behaviour is conceived as being directed towards the control of perceptions. Thus, PCT has some of the characteristics that would be desirable in an integrated theory covering psychological and situational variables. Dick Pew, who also writes in these proceedings, suggests that Neisser's theory of the perceptual cycle (Neisser, 1976) has some of the characteristics required to deal with the interaction between the perceiver and the environment (Tenney, Adams, Pew, Huggins, & Rodgers, 1992).

Most theories of cognitive psychology have evolved through testing in controlled, laboratory settings, using single rather than multiple tasks, in static rather than dynamic task environments. The consequences are that findings can be difficult to generalise to real world problems. This has been found to be a concern in applying the theory of decision-making to real world decisions (Klein, Orsanu, & Calderwood, 1992). However, there is a increasing trend towards testing and developing cognitive theory applicable to naturalistic, "every-day" settings.

One important consequence of relevant psychological theory is the ability it provides to create computational models for testing predictions of human performance in simulated task environments. Human performance models are being used in aircrew systems design to prototype tasks, and to predict operator performance with increasingly complex, dynamic task and system variables. On the UK RN Merlin helicopter programme, for example, a relatively simple model of attentional demand has been used to identify potential operator task loadings and workload "bottle-necks" in comparison with a "situational" model (mission "story line") of predicted task demands. Also, dynamic task network simulation is being used on the Merlin project to predict the consequences of operator decision errors on mission effectiveness (MacLeod , Biggen, Romans, & Kirbyet, 1993). Under the auspices of The Technical Cooperation Programme (TTCP), in collaboration with DCIEM, the RAF IAM is

conducting research to investigate ability to predict operator SA using dynamic task network simulation and a relatively simple, attentional demand model of human performance. The question here will be how well SA can be predicted from a theory of attentional demand without taking into account knowledge variables.

Another important consequence of psychological theory is that computational models can be embedded in aircraft software to provide real-time predictions of human performance for adaptive aiding. The USAF/DARPA Pilot's Associate programme proposes a Pilot-Vehicle Interface with pilot intent inferencing and an adaptive aiding concept. Predictions of pilot requirements are based on embedded models of situational variables, human performance and human error, and on comparison of monitored situational variables and pilot actions with scripts, plans and goals (Andes 1987).

The Findings?

The question here is, "do the theoretical predictions fit experience and the experimental data?" There are limited data from which to draw conclusions. One of the major problems in testing predictions is that it requires measurement of SA. This raises the issue of the validity of measuring an inferred cognitive construct. The requirement for measurement of SA has received considerable attention. There are a number of approaches, involving objective and subjective measures, and there has been considerable debate about their appropriateness. The issues have been reviewed in some detail elsewhere (Fracker & Vidulich 1991). Objective measures would be preferable, but, performance-based measures of SA require assumptions to be made about causal relationships. The lack of an integrated theory limits the inferences that can be drawn about underlying cognitive processes.

The approach that we have taken to SA measurement at the RAF IAM has been to develop subjective measures. This approach was taken to complement rather than to substitute for objective measures. We were encouraged by the apparent high utility of subjective workload measures, which has been attributed to their ease of implementation, low intrusiveness, and good operator acceptance. The method we have developed is called the Situational Awareness Rating Technique or SART (Taylor 1989). SART is based on the aircrew constructs associated with SA referred to earlier. With SART, subjective ratings are obtained for the aircrew constructs for attentional demand, attentional supply, and understanding. Considerable attention has been given to establishing the validity, sensitivity and diagnostic power of the rating scales. SART is available from the IAM in 3, 10, and 14 dimension forms, implemented using paper and pencil, and in software versions.

Criticism has been raised over the "calibration problem" with subjective SA ratings. The ratings must be made relative to some notion of what is not known. SART partly addresses this criticism by requiring ratings of attentional demand and supply. The difference between the estimates of demand and supply provide a crude estimate of what is unknown. SART has been shown to have utility in the assessment of human performance in a variety of skill, rule, and knowledge-based tasks, including tracking and monitoring aircraft HUD flight parameters, unusual aircraft attitude recovery, aircraft warnings comprehension, and aircraft flight simulation (Selcon & Taylor 1990; Taylor & Selcon 1990; Selcon, Taylor, & Koritsas, 1991). Because of poor sensitivity and diagnosticity, we have found little utility in uni-

dimensional ratings of SA. However, we are currently investigating the possibility of deriving a single measure from the 3-D SART ratings of attentional demand, supply, and understanding. On the basis of *a priori* theoretical considerations, and as an initial working hypothesis, we have proposed that SA might be calculated by combining the ratings on the three SART dimensions by using the following formula:

SA(c) = U - (D-S)
where:
SA(c) is calculated Situational Awareness
U is rated Understanding
D is rated Attentional Demand
S is rated Attentional Supply

Initial data show some evidence that the SA(c) scores reflect performance data on a bi-modal warnings comprehension task (Selcon, Taylor, & Shadrake, 1992). Other SART data, reported in these proceedings by Dr. Mike Vidulich of the USAF Armstrong Laboratory, show a relationship between SA(c) scores and the effects of display variables on performance on a flight simulation task (i.e. STORM). Further research is needed to clarify the validity of treating SA as a uni-dimensional scalar concept, rather than treating the separate 3-D SART scores as individual vector quantities.

Conclusions

There seems to be a problem with SA in advanced systems, particularly with high levels of task automation. This difficulty is in part due to the lack of an integrated theory which accounts for both the environmental and psychological factors involved in situational awareness. The psychological theories proposed have not generated sufficient predictions and proven hypotheses to be confident that the problems are well understood and that solutions are at hand. Measurement is needed to test predictions from theory. However, the measurement of an inferred cognitive construct such as SA raises complex validity and methodological issues. Some progress has been made in developing subjective measures to supplement objective data, but so far there is limited evidence on which to draw firm conclusions about the adequacy of the theory and the best solutions to the problem.

References

Andes R.C. (1987). Adaptive aiding in complex systems: an implementation, In, *Proceedings of the IEEE 1987 Conference on Systems, Man, and Cybernetics.*
Bundeson C. (1990). A theory of visual attention. *Psychological Review, 97,* 523-547.
Chechile R.A., Egglestone R.G.,Fleishman R.N., & Sasseville A.M. (1989).*Human Factors, 31,* 31-43.

Endsley, M.R. (1988). Design and evaluation for situation awareness enhancement. In, *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97-101). Santa Monica,CA: The Human Factors Society.

Fracker M. L. (1988). A theory of situation awareness: implications for measuring situation awareness. In, *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 102-106). Santa Monica, CA: The Human Factors Society.

Fracker M.L. & Vidulich M.A. (1991). Measurement of situation awareness: a brief review. In, Y. Queinnec & F. Daniellou (Eds.), *Designing for Everyone, Proceedings of the 11th Congress of the International Ergonomics Association,* (pp. 795-797). London: Taylor & Francis.

Freyd J.J. (1987). Dynamic mental representations. *Psychological Review, 94,* 427-438.

Howard, I. P. (1982). *Human visual orientation*. Chichester: Wiley.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge: University Press.

Rasmussen, J. (1986). *Information processing and human-machine interaction.* Amsterdam: Elselvier.

Kinchla R.A. (1980), The measurement of attention. In R.S. Nickerson (Ed.), *Attention and Performance VIII* (pp. 213-238). Hillsdale, NJ : Erlbaum.

Klein, G.A., Orsanu, J., & Calderwood, R. (1992). *Decision making in action: models and methods.* Norwood, NJ: Ablex.

MacLeod, I.S., Biggen, K., Romans, J.& Kirby K. (1993). Predictive workload analysis- RN EH101 helicopter. In, E.J. Lovesey (Ed.), *Contemporary Ergonomics 1993.* London: Taylor & Francis (In press).

Neisser, U. (1976). *Cognition and reality: principles and implications of cognitive psychology.* San Francisco: Freeman.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. Amsterdam: Elselvier.

Sarter N.B. & Woods D.D. (1991). Situation awareness: A critical but ill-defined problem. *International. Journal of Aviation Psychology, 1,* 45-57.

Schank R.C. & Abelson R. (1977). *Scripts, plans ,goals and understanding.* Hillsdale NJ: Erlbaum.

Selcon S.J. & Taylor R.M. (1990). Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design. In, *AGARD-CP-478* (pp. 5-1to 8). Neuilly-sur-Seine: NATO AGARD.

Selcon S.J. & Taylor R.M. & Koritsas E. (1991). Workload or situational awareness?: NASA TLX versus SART for aerospace systems design. In, *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 62-66). Santa Monica, CA: The Human Factors Society.

Selcon S.J. & Taylor R.M. & Shadrake R. (1992). Multi-modal cockpit warnings: pictures, words or both.In, *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 57-61). Santa Monica, CA: The Human Factors Society.

Taylor M.M. (1988). Layered protocols for computer-human dialogue.I.: Principles, *International Journal of Man-Machine Studies, 28* (219-257).

Taylor R.M. (1987). Workload management in military cockpits: Present insights. In, *Recent Advances in Cockpit Aids for Military Operations*. London: Royal Aeronautical Society, .

Taylor R.M. (1989). *Situation awareness: Aircrew constructs for situation awareness,* IAM Report No 670. Farnborough, Hants: RAF Institute of Aviation Medicine.

Taylor R.M. & Selcon, S.J. (1990). Cognitive quality and situational awareness with advanced aircraft attitude displays. In, *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 26-30). Santa Monica, CA: The Human Factors Society.

Tenney, Y.J., Adams, M.J., Pew, R.W., Huggins, A.W.F. & Rodgers, W.H. (1992). A principled approach to the measurement of situation awareness in commercial aviation. NASA Contractor Report 4451. Hampton, VA: NASA Langley Research Center.

Triesman A. & Gelade G. (1980). A feature integration theory of attention, *Cognitive Psychology, 12*, 97-136.

Yantis S. & Johnson D.N. (1990). Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 812-815.

Situation Awareness in Dynamic Human Decision Making: Measurement

Mica R. Endsley

Texas Tech University

Methodologies for the empirical measurement of situation awareness are reviewed, including a discussion of the advantages and disadvantages of each method and the potential limitations of the measures from a theoretical and practical viewpoint. Two studies are presented which investigate questions of validity and intrusiveness regarding a query based technique. This technique requires that a simulation of the operational tasks be momentarily interrupted in order to query operators on their situation awareness. The results of the two studies indicate that the query technique is not intrusive on normal subject behavior during the trial and that the technique does not suffer from limitations of human memory, providing an indication of empirical validity. The results of other validity studies regarding the technique are discussed briefly along with the use of this technique for measuring situation awareness in varied settings.

Introduction

Situation awareness (SA), an operator's internal model of the surrounding world, is a key ingredient for effective decision making in a dynamic environment. Operators of dynamic systems must ascertain the current status and dynamics of their systems and other relevant elements in the environment in order to determine the best course of action to take at any point in time. Without this knowledge, most operators will not be able to function satisfactorily. This is true for many systems, including: aircraft; air traffic control; large systems, such as flexible manufacturing systems, refineries, and nuclear power plants; strategic systems such as fire fighting units, certain police units and military command centers; and for many daily activities such as driving.

Situation awareness will be defined formally as *"the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"* (Endsley, 1987,1988b). The first step in achieving SA is for the operator to perceive the appropriate elements in the environment (Level 1 SA). But SA involves far more than simple perception. SA also involves *understanding* that information – integrating the various separate elements to form a whole

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

picture (gestalt) and comprehending the significance of this information in light of operator goals (Level 2 SA). Finally, those with superior SA will be able to project the future behavior of elements based on that understanding (Level 3 SA). A well developed internal system model is usually required for this. As decision selection and performance flows directly from situational understanding, often in an automatic fashion, the formation of a correct and complete situation model (or SA) is critical.

In many environments, the achievement of good SA is a highly taxing process due to the complexity of systems, the presence of large quantities of information which may change rapidly, and high workload among other issues. While training, expertise and individual capabilities are important factors determining the degree to which a given individual may achieve SA in these environments, the design of the human machine interface (HMI) will also have a major and critical role in this process.

In light of this, the enhancement of SA has developed as a major goal for human factors practitioners who are developing control and display designs, automation concepts or training programs in a variety of fields. In the aircraft industry, for example, there have been a plethora of solutions suggested for improving SA. Press (1986) states that SA can be increased through improvements in selection, aircraft technology, tactics and training. Other authors have advocated improved sensor capabilities (Stiles and Pearson, 1986), improved, integrated eyes-out controls and displays (Dornheim, 1986; Person and Steinmetz, 1981), intelligent systems which will integrate, prioritize, filter, and communicate information to the pilot based on the situation (Morishige and Retelle, 1985), and systems to reduce workload through the automation of certain pilot tasks (AirForce, 1982; Chambers and Nagel, 1985; Jurgensen and Feldman, 1985; O'Shannon, 1986).

To achieve real improvements in operator SA, in aircraft or other systems, it is necessary to determine which of such ideas have merit and which might perhaps have unforeseen negative consequences. To resolve this issue, each concept's effect on operator SA must be evaluated. The process of developing systems which will provide human operators in complex systems with SA will greatly benefit from the ability to evaluate the impact of alternate design concepts on operator SA. Only in this way can a concept's utility be established.

A measure of SA would also be quite useful for expanding the knowledge-base on situation awareness. With a means of measuring SA available, it should be possible to (a) examine sources of individual differences in SA, (b) conduct investigations of mental models, (c) more adequately assess decision making through a better understanding of the inputs to that process, and (d) assess the relative contributions of SA components.

Measurement Techniques

In any design process, the use of iterative, manned simulation testing to evaluate competing design concepts is needed in order to detect problems with given designs and to ascertain the best of competing concepts (often in the form of rapid prototyping or part-task simulation). Several different methods have been attempted or can be considered for the measurement of SA during this type of testing. In addition, many of these techniques could be used to evaluate the SA of operators working with actual systems. For the most part, efforts at measuring SA

thus far have been concentrated in the aircraft environment, however, most of the techniques are equally applicable to other types of systems.

Physiological Techniques

Ideally, it would be desirable to install a window on the operator's mind and observe an exact picture of what is known at all times. Unfortunately, no device currently exists for doing this. Some less intrusive physiological measures might be proposed, however. P300 and other EEG measurements, for instance, have shown promise for determining whether information is registered cognitively. These techniques will allow researchers to determine if elements in the environment are perceived and processed by subjects, but do not allow a determination of how much information remains in memory, if the information is registered correctly in the mind, or what comprehension the subject has of those elements. Similarly, eye-tracking devices appear to fall short for the same reasons. Furthermore, they will not tell which elements in the periphery of the subject's vision are observed, or if the subject has even processed what was looked at. Therefore, known physiological techniques, while providing useful, objective data, are not very promising for SA measurement.

Performance Measures

In general, performance measures provide the advantage of being objective and are usually non-intrusive. Computers for conducting system simulations can be programmed to record specified performance data automatically, making the required data relatively easy to collect. Several limitations exist in using performance data to infer SA, however.

Global measures. Global measures of performance suffer from problems of diagnosticity and sensitivity. While overall operator/system performance is always a useful bottom-line criterion for evaluating competing concepts, if that is the only criterion, important system differences can be masked. Performance measures give only the end result of a long string of cognitive processes, providing little information about why poor performance may have occurred in a given situation (if it can be detected reliably during testing at all). Poor performance could be due to a lack of information, poor sampling strategies, improper integration or projection, high workload, poor decision making, or action errors among other factors, many of which are not SA problems and indicate completely different solutions.

In addition, in many cases overall system performance is not a terribly useful criterion as it will be masked by other factors. In a tactical aircraft environment, for instance, much of pilot performance is, by nature, highly variable and subject to the influence of many other factors besides SA. A new system may provide the pilot with better SA, but in evaluation testing this fact can be easily masked by excessive workloads, the intentional use of varied tactics or poor decision making if overall mission performance is used as the only dependent measure. It would be desirable, therefore, to measure SA more directly.

External task measures. One type of performance based measure which has been suggested involves artificially changing certain information or removing certain pieces of information from operator displays and then measuring the amount of time required for the operator to react to this event. Aside from the fact that such a manipulation is heavily

intrusive and requires the subject to undertake new tasks involved with discovering what happened to the changed or missing data while attempting to maintain satisfactory performance on other tasks, this technique may provide highly misleading results. It assumes that an operator will act in a certain way, given a change in or disappearance of certain data, when, in fact, operators often employ work-around schemes to function under just such circumstances. For instance, if a displayed aircraft suddenly disappears, the operator may assume (a) malfunctioning equipment, (b) the aircraft was destroyed or landed, or (c) the aircraft's emitting equipment was turned off, rendering it more difficult to detect (a not infrequent occurrence). In any case, the operator may chose to (a) ignore the disappearance, assuming it will come back on the next several sweeps of the radar, (b) worry about it, but not say anything, or (c) put off dealing with the disappearance until other tasks are complete. Any of these actions will yield highly misleading results for the experimenter who expects subject SA to be reflected by the operator's behavior.

The bottom-line is not only that such assumptions are invalid, but also that it would seem prudent to avoid any technique which fundamentally alters the subjects' ongoing tasks, as situation awareness itself can be easily altered in the process. Anytime one artificially alters the realism of the simulation, it could fundamentally effect the way the operator conceptualizes the underlying information (see Manktelow and Jones, 1987), thus altering both SA and decision making. In addition, such manipulation would certainly interfere with any concurrent workload or performance measurement undertaken during the trial.

Imbedded task measures. Some information about SA can be determined from examining performance on specific operator subtasks that are of interest. For example, when evaluating an altitude display, deviations from proscribed altitude levels or time to reach a certain altitude can be measured. This type of detailed performance measure can provide some inferences regarding the amount of SA about a specific parameter that is provided by a certain display. Such measures will be more meaningful than global performance measures and will not suffer from the same problems of intrusiveness as external task measures. While finite task measures may readily present themselves for evaluating certain kinds of systems, for others, however, determining appropriate measures may be more difficult. An expert system, for instance, may influence many factors in a global, not readily predicted manner.

The major limitation of this approach stems from the interactive nature of situation awareness sub-components. A new system to provide SA on one factor may simultaneously reduce SA on another, not measured, factor. Endsley (1989b), for instance, found that a three-dimensional display increased SA on the z-dimension (altitude), only at the expense of SA on the x and y dimensions (range and azimuth). Similarly, Fracker (1989) found that SA was increased on certain objects in a display at the expense of others. In addition, it is quite easy for subjects to bias their attention to a single issue which is under evaluation in a particular study (e.g. altitude) if they figure out the purpose of the stu[illegible]verall, as improved SA in one area may easily result in decreased SA in others, re[illegible]g on the measurement of performance on specific parameters can yield misleading results.

What researchers really need to know is: how much SA do operators have when taxed with all of the multiple, competing demands upon their attention that occur during system operations? For this reason, a global measure of SA which simultaneously depicts SA across the many elements of interest is desirable. To improve situation awareness, designers need to be able to evaluate the impact of design concepts on operator SA in its entirety.

Subjective Techniques

Self- rating. One very simple technique that has been used occasionally is to ask operators to subjectively rate their own SA (e.g. on a 1 to 10 scale). The AMRAAM OUE study (1982) used this method. Pilot (and overall flight) SA was subjectively rated by the participants and by a trained observer. The main advantages of subjective estimation techniques are that they are inexpensive and easy to use. In general, however, the subjective self-rating of SA has several limitations.

1. If the ratings are collected during a simulation trial, the operators' ability to estimate their own SA will be quite limited since they do not know what is *really* happening in the environment (they only have their *perceptions* of that reality). Operators may know when they do not have a clue as to what is going on, but will probably not know if their knowledge is incomplete or inaccurate.

2. If operators are asked to subjectively evaluate SA in a post-trial debriefing session, the rating may also be highly tainted by the outcome of the trial. When performance is favorable, whether through good SA or good luck, an operator will most likely report good SA, and vice-versa. A re-evaluation of the AMRAAM OUE study by Venturino, Hamilton and Dvorchak (1989) found just that. Post-trial subjective SA ratings were highly correlated with performance.

3. When ratings are gathered after the mission, operators will probably be inclined to rationalize and over generalize about their SA, as has been shown to be the case when information about mental processes is collected after the fact (Nisbett and Wilson, 1977). Thus, detailed information will be lost or misconstrued.

So, what do such estimates actually measure? I would speculate that subjective self-ratings of SA most likely convey a measure of the subjects' confidence level regarding that SA. That is, how comfortable they feel about their SA. Subjects who know they have incomplete knowledge or understanding would subjectively rate their SA as low. Subjects whose knowledge may not be any greater, but who do not subjectively have the same concerns about how much is not known, would rate their SA higher. In other words, ignorance may be bliss.

Several efforts have been made to develop more rigorous subjective measures of SA. Taylor (1989) has developed the Situation Awareness Rating Technique (SART) which allows operators to rate a system design on the amount of demand on attentional resources, supply of attentional resources and understanding of the situation provided. As such, it considers operators' perceived workload (supply and demand on attentional resources) in addition to their perceived understanding of the situation. While SART has been shown to be correlated with performance measures (Selcon and Taylor, 1989), it is unclear whether this is due to the workload or the understanding components. The other limitations of subjective SA techniques also apply.

In a new application, the Subjective Workload Dominance (SWORD) metric (Vidulich, 1989) has been applied as a subjective rating tool for SA (Hughes, Hassoun, and Ward, 1990). SWORD allows subjects to make pairwise comparative ratings of competing design concepts along a continuum expressing the degree to which one concept entails less workload than the other. The resultant preferences are then combined using the analytic hierarchy process (AHP) technique to provide a linear ordering of the design concepts. In the Hughes et al. study, this

technique was modified to allow pilots to rate the degree to which presented display concepts provided SA instead of workload. Not surprisingly, the display that was subjectively rated as providing the best SA using SWORD was the display that pilots expressed a strong preference for. It is difficult to ascertain whether subjective preference led to higher SWORD ratings, or vice-versa. More research is needed to determine the locus of such ratings.

Observer-rating. A second type of subjective rating involves using independent, knowledgeable observers to rate the quality of a subject's SA. While a trained observer might have more information than the operator about what is really happening in a simulation (through perfect knowledge gleaned from the simulation computer), the observer would have very limited knowledge about what the operator's concept of the situation is. The only information about the operator's perception of the situation would come from operator actions and imbedded or elicited verbalizations by the operator (e.g., from voice transmissions during the course of the task or from verbal protocols explicitly requested by the experimenter). While this knowledge can be very useful diagnostically, to determine overt errors in SA (stated misperceptions or lack of knowledge) for instance, it cannot be said to be in any way a complete representation of that knowledge. The operator may, and in all likelihood will, have a much greater store of information held internally which is not verbalized. For instance, a pilot may discuss his efforts to ascertain the identity of a certain aircraft, but his knowledge about his ownship system status, heading, other aircraft, etc., might not be mentioned at all. An outside observer has no way of knowing whether the pilot is aware of these variables, but is not discussing them because they are not of immediate concern, or whether the pilot has succumbed to attentional narrowing and has no idea of the real value of these parameters. As such, the use of outside observers to rate SA is also limited.

A variation on this theme is to use a confederate who acts as an associate to the operator (e.g., another crew member or air traffic controller in the case of a commercial aircraft), and requests certain information from the operator to encourage further verbalization, as has been suggested by Sarter and Woods (1991). In addition to the same limitations encumbering observer ratings, this technique may also serve to *alter* SA in the experimental setting by artificially directing the subject's attention to certain parameters. As the distribution of the subject's attention across the elements in the environment largely determines SA, this method probably does not provide an unbiased assessment of operator SA.

Questionnaires

In general, questionnaires allow for detailed information about subject SA to be collected which can then be evaluated against *reality,* thus providing an objective assessment of operator SA on a detailed level. This type of assessment is a more direct measure of SA (i.e., it taps into the operator's perceptions rather than infers them) and does not require subjects or observers to make judgments about situational knowledge on the basis of incomplete information, as subjective assessments do. In a review of such questionnaires, Herrmann (1984) concludes that when information reported in this manner can be evaluated on the basis of objective knowledge, they have been found to have good validity. Several methods of administration are possible.

Post-test. A detailed questionnaire can be administered after the completion of each simulated trial. This allows ample time for subjects to respond to a lengthy and detailed list of

questions about their SA during the trial, providing needed information about subject perceptions. Unfortunately, people in general are not good at reporting detailed information about past mental events, even recent ones. There is a tendency to over generalize, over summarize and over rationalize. Recall will be stilted by the amount of time and by intervening events which occur between the activities of interest and the administration of the questionnaire (Nisbett and Wilson, 1977). Earlier misperceptions can be quickly forgotten as the real picture unfolds itself during the course of events. Therefore, a post-test questionnaire will really only reliably capture the subject's SA at the very end of the trial.

On-line. One way of overcoming this deficiency is to ask operators about their SA *while* they are carrying out their simulated tasks. Unfortunately, this too has several drawbacks. First of all, in many situations of interest, the subject will be under very heavy workload, precluding the answering of additional questions. Such questions would also constitute a form of secondary task loading that may alter performance on the main task of operating the system. Furthermore, the questions asked could cue the subject to attend to the requested information on the displays, thus altering the operator's *true SA*. An assessment of time to answer as an indicator of SA is also faulty, as various strategies may be employed by subjects who are time sharing between the dual tasks of operating the system and answering the questions. Overall, this method will be highly intrusive on the primary task of system operation.

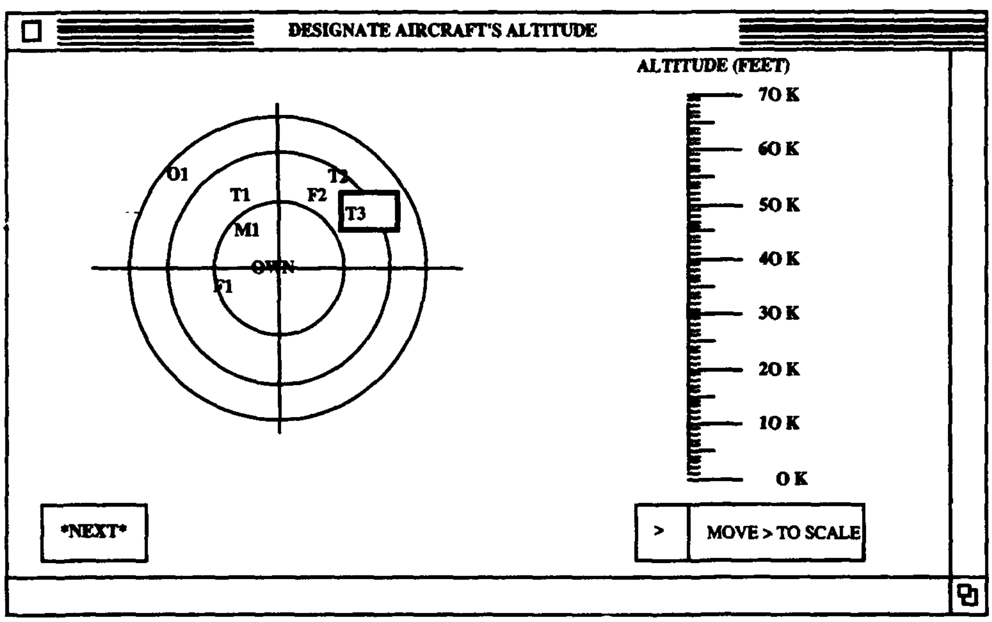


Figure 1. Sample SAGAT Query

Freeze Technique. To overcome the limitations of reporting on SA after the fact, several authors have used a technique wherein the simulation is frozen at randomly selected times

and subjects are queried as to their perceptions of the situation at that time (Endsley, 1987,1988a, 1988b; Fracker, 1989; Marshak, Kuperman, Ramsey, and Wilson, 1987). With this technique the system displays are blanked and the simulation is suspended while subjects quickly answer questions about their current perceptions of the situation. Thus, SA data can be collected immediately, reducing the problems incurred when collecting data after the fact, but not incurring the problems of on-line questioning. Subject perceptions are then compared to the real situation based on simulation computer data bases to provide an objective measure of SA.

In a study evaluating competing aircraft display concepts, Marshak, et al. (1987) queried subjects about navigation, threats and topography using this technique. The resulting answers were converted to an absolute percent error score for each question, allowing scores across displays to be compared along these dimensions. Fracker (1989), in another aircraft oriented study, requested information on aircraft location and identity for specific indicated aircraft in the simulation. In both studies, a measure of SA on only select parameters was obtained.

The Situation Awareness Global Assessment Technique (SAGAT) is a global tool developed to assess SA across all of its elements based on a comprehensive assessment of operator SA requirements (Endsley, 1987,1988a, 1988b). As a global measure, SAGAT includes queries about all operator SA requirements, including Level 1 (perception of data), Level 2 (comprehension of meaning) and Level 3 (projection of the near future) components. This includes a consideration of system functioning and status as well as relevant features of the external environment.

Computerized versions of SAGAT have been developed for air-to-air tactical aircraft (Endsley, 1987, 1990b) and advanced bomber aircraft (Endsley, 1989a) which allow queries to be administered rapidly and SA data to be collected for analysis. The tool features an easy to use query format that was designed to be as compatible as possible with subject knowledge representations. An example of a SAGAT query is shown in Figure 1. SAGAT's basic technique is generic and is applicable to other types of systems once a delineation of SA requirements has been made.

This approach has an advantage over probes which only cover a limited number of SA items in that subjects cannot prepare for the queries in advance since they could be queried over almost every aspect of the situation to which they would normally attend. Therefore, the chance of biasing subject attention to specific items is minimized. SAGAT has thus far been used to evaluate competing aircraft avionics concepts (Northrop, 1988), various display hardware and software concepts (Bolstad and Endsley, 1990; Endsley, 1989b), the relationship between SA and workload (Endsley, in preparation) and SA and performance (Endsley, 1990a), and to investigate individual differences in SA (Endsley and Bolstad, in preparation).

The collection of SA data in this manner provides several advantages. It overcomes the problems incurred when collecting data after the fact, yet minimizes biasing of subject SA due to secondary task loading or artificially cueing the subject's attention. Secondly, it provides a *direct* measure of SA which can be *objectively* collected and *objectively* evaluated. Third, the use of random sampling (both of stop times and query presentation) provides unbiased estimates of SA, thus allowing SA scores to be easily compared statistically across trials, subjects and systems.

The primary disadvantage of this technique involves the temporary halt in the simulation. Two major issues must be addressed.

1. Can subjects report their SA using this technique, and if so, for how long?

2. Does temporarily freezing the simulation result in any change in subject behavior? That is, is it intrusive in such a manner that performance will be altered? Two experiments were conducted to address these concerns.

Experiment 1

In determining whether SA information will be reportable via the SAGAT methodology, several possibilities must be considered.

1. Data may be processed by subjects in short term memory (STM), never reaching long term memory (LTM). If a sequential information processing model is used, then it is possible that information might enter into STM and never be stored in LTM where it would be available for retrieval during questioning. In this case, information would not be available during any SAGAT query sessions which exceeded the STM storage limitations (approximately 30 seconds with no rehearsal).

   There is a good deal of evidence, however, that STM may not precede LTM, but merely be an activated subset of LTM (Cowan, 1988; Morton, 1969; Norman, 1968). According to this type of model, information proceeds directly from sensory memory to LTM, which is necessary for pattern recognition and coding. Only those portions of the environment which are salient are then highlighted in STM (either through focalized attention or automatic activation). This type of model would predict that SA information which has been perceived and/or further processed by the operator would exist in LTM stores and thus be available for recall during SAGAT querying which exceeds 30 seconds.

2. The data may be processed in a highly automated fashion, and thus not be in the subject's awareness. Expert behavior can function in an automated processing/action sequence in some cases. Several authors have found that even when effortful processing is not used, however, the information is retained in LTM and is capable of affecting subject responses (Jacoby and Dallas, 1981; Kellog, 1980; Tulving, 1985). The type of questions used in SAGAT, providing cued-recall and categorical or scalar responses, should be receptive to retrieval of this type of information.

3. The information may be in LTM, but not be easily recalled by the subjects. Evidence suggests that when effortful processing and awareness are used during the storage process, recall is enhanced (Cowan, 1988). SA, composed of highly relevant, attended to and processed information, should be most receptive to recall. In addition, the SAGAT battery, requiring categorical or scalar responses, is a cued recall task, as opposed to total recall, thus aiding retrieval. Under conditions of SAGAT testing, the subjects are aware that they may be asked to report their SA at any time. This too may aid in the storage and retrieval process. Since the SAGAT battery is administered immediately after the freeze in the simulation, no time for memory decay or competing event interference is allowed. Thus, the conditions

should be optimized for the retrieval of the SA information. While it cannot be said conclusively that all of the subject's SA can be reflected in this manner, the vast majority should be reportable via the SAGAT technique.

To further investigate this matter, a study was conducted to specifically determine how long after a freeze in the simulation SA information could be obtained. It was expected that if collection of SA data via this technique was memory limited, this would be evidenced by an increase in errors on SAGAT queries occurring around 30 to 60 seconds after the freeze time due to short term memory restrictions. (While this would not preclude use of the technique, it would limit its use by restricting the number of questions that could be asked at a particular stop in the simulation.) A simulation of a tactical aircraft task was used for the study.

Procedure

A set of air-to-air engagements was conducted at the Northrop Aircraft Division in their real-time, manned, multi-engagement simulator facility. A fighter sweep mission with a two (blue team) versus four (red team) force ratio was used for the trials. The objective of the blue team was to penetrate red territory, maximizing the kills of red fighters while maintaining a high degree of survivability. The red team was directed to fly around their assigned combat air patrol (CAP) points until a blue target was detected in red airspace. They were then allowed to leave their CAP point to defend against the blue team. In all cases, specific tactics were at the discretion of the individual pilot teams.

A total of fifteen trials was completed by two teams of six subjects. At a random point in each trial, the simulator was frozen and SAGAT data immediately collected from all six participants. At a given stop, each of the queries was presented once, in a random order. As this order was different at each stop, each query was presented at a variety times after the stop across subjects and trials. As all subjects answered all the queries in the SAGAT battery at each stop, approximately 90 data points per query were obtained. Each point provided an index of performance corresponding to a different amount of elapsed time after the stop. After all subjects had completed the SAGAT battery at a given stop, a new trial was begun.

Prior to conducting the study, all subjects were trained on the use of the simulator, the displays, aircraft handling qualities, and SAGAT. In addition to three instructional training sessions on using SAGAT, each subject participated in 18 practice trials in which SAGAT was administered. (Most subjects also had received a substantial amount of training in the simulator in the past.) Thus, the subjects were well trained prior to testing.

Facilities. Northrop's Integrated Simulation and Systems Laboratory (ISSL) was used for the test. ISSL is a high-fidelity, real-time, interactive, man-in-the-loop facility which incorporates simulation technologies that are used in the design, development, and evaluation of tactical aircraft and associated weapon systems. Aircraft control systems, avionics systems, weapons systems, crew stations and air vehicle performance are all modeled to provide aircraft design variations for testing in multiple engagement simulation. ISSL incorporates a Gould mainframe computer which controls simulations and drives Silicon Graphics generated high-resolution color graphics displays. This test used six manned stations, each configured to represent hypothetical future generation aircraft. (Hypothetical performance characteristics, weapons, and avionics capabilities were used in order to keep the simulation at a unclassified level.) Each control station includes a simulated head-up display,

a tactical situation display, radar and system controls operated by a touch screen or stick and throttle control switches. A realistic stick and throttle provide primary flight control.

Subjects. Twelve subjects participated in this test. The subjects were all experienced former military fighter pilots currently employed by Northrop. The mean subject age was 48.16 years (range of 32 to 68). They had an average of 3310 hours (range of 1500 to 6500) and an average of 15.5 years (range of 7 to 26) of military flight experience. Seven of the twelve subjects had combat experience.

Hypotheses. H_0: There is no difference in subject performance on SAGAT queries as a function of the amount of time after the stop that the query is presented.

H_1: There is a difference in subject performance on SAGAT queries as a function of the amount of time after the stop that the query is presented.

Results

Each of the subjects' answers to the SAGAT queries were compared to actual values, as collected by the simulator computer at the time of each stop. Twenty-six queries were included in the SAGAT battery at the time of this test. Of those, eleven could not be evaluated because the appropriate simulator data was not available and five could only be evaluated by subjective means. Answers to the ten remaining queries were evaluated. These include: ownship heading, ownship location, aircraft heading, aircraft detections, aircraft airspeed, aircraft weapon selection, aircraft Gs, aircraft fuel level, aircraft weapon quantity, and aircraft altitude.

An error score (SAGAT query answer - actual value) was computed for each response. Absolute error scores for each query (across all subjects and trials) were plotted against the amount of time after the stop in the simulation that the query was asked, and a regression calculated.

None of the regressions computed for each of the ten queries was significantly different than zero ($\alpha = .05$), indicating that subjects were neither more nor less prone to response error as the amount of time between the simulator freeze and the presentation of the query increased. A plot of the regression for altitude error, shown in Figure 2, reveals little or no increase in error over time. (Plots of the regressions for the other variables appeared quite similar.)

Discussion

Based on this data, it would appear that subjects were able to provide information on their SA about a particular situation for up to five or six minutes under these conditions. The fact that all ten of the queries produced very flat regressions lends extra weight to this conclusion.

Two explanations can be offered for these findings. First, this study investigated expert subjects' knowledge of information which was extremely important to task performance during a realistic simulation of those tasks. Most laboratory studies which predict fairly rapid decay times (approximately 30 seconds for short-term memory) typically employ the use of stimuli which have little or no inherent meaning to the subject (nonsense words or pictures).

Studies have found that the storage and utilization of relevant information may be quite different than that of irrelevant information (Chase and Simon, 1973).

Second, the results indicate that the SA information was obtainable from long-term memory stores. If schema, or other mechanisms, are used to organize SA information (as opposed to working memory processes only), then that information will be resident in long-term memory (LTM). Many of the ten items analyzed can be considered Level 1 SA components. The fact that this lower level information was resident in LTM indicates that either (a) the inputs to higher level processing were retained as well as the outputs, or (b) the Level 1 components were retained as important pieces of pilot SA in their own right and are significant components of LTM schema (e.g., target altitude itself is important to know and not just the implications of target altitude). Both of these explanations may be correct. These findings generally support the predictions of a processing model in which information passes into LTM storage before being highlighted in STM.

As a caveat, it should be noted that the subjects were actively working with their SA knowledge by answering the SAGAT queries for the entire period that the simulation was stopped. No intervening period of waiting nor any competing activity was introduced prior to administering any SAGAT query. Subject knowledge of SA information may be interfered with if time delays or other activities (particularly continued operational tasks) are imposed before SAGAT is administered. The major implication of these results is that, under these conditions, SA data is readily obtainable through the SAGAT technique for a considerable period of time, up to five or six minutes, after a stop in the simulation.

Experiment 2

A second study was initiated to address the issue of possible intrusiveness. It is necessary to determine whether temporarily freezing the simulation results in any change in subject behavior. While not directly ruling out the use of this technique, if subject behavior is altered by such a freeze, certain limitations would be indicated. (One might wish to not resume the trial after the freeze, for instance, starting a new trial instead.) The possibility of intrusiveness was tested by evaluating the effect of stopping the simulator on subsequent subject performance.

Procedure

A set of air-to-air engagements was conducted of a fighter sweep mission with a two (blue team) versus four (red team) force ratio. The training, instructions and pilot mission objectives were identical to those used in Experiment One. In this study, however, the trial was resumed after freeze following a specified period for collecting SAGAT data, and was continued until specified criteria for completion of the mission were met. The subjects completed as many queries as they could during each stop. The queries were presented in a random order.

Five teams of six subjects completed a full test matrix. The independent variables were duration of the stops (one-half, one or two minutes) and the frequency of stops (one, two, or

three times during the trial). Each team participated twice in each of these nine conditions. (In any given trial, multiple stops were of the same duration.) Each team also completed six trials in which no stops occurred as a control condition. Therefore, a total of 30 trials were conducted for each duration of stop condition (one-half, one or two minutes) and a total of 30 trials were conducted for each frequency of stop condition (one, two, or three). These conditions could be compared to 30 trials in which no stops occurred. Conditions were administered in a random order, blocked by team. Pilot performance was collected as the dependent measure.

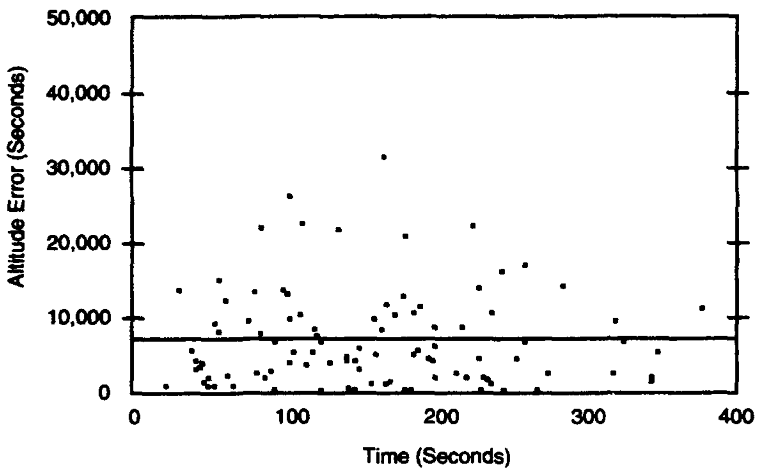


Figure 2. Altitude error by time until query presentation.

Facilities. This study used the same manned, multiple-engagement simulator as Experiment One.

Subjects. Twenty-five subjects participated in this test. (Five of the subjects participated on more than one team.) The subjects were all experienced former military fighter pilots employed by Northrop. The mean subject age was 45.16 years (range of 32 to 68). They had an average of 3582 hours (range of 975 to 7045) and an average of 16.9 years (range of 6 to 27) of military flight experience. Fourteen of the 25 subjects had combat experience.

Hypotheses. H_0: There is no difference in pilot performance between trials in which there are stops to collect SAGAT data and trials in which there are no stops.

H_1: There is a difference in pilot performance between trials in which there are stops to collect SAGAT data and trials in which there are no stops.

Results

Pilot performance under each of the conditions was analyzed. Pilot performance measures included the number of blue team kills (red team losses) and the number of blue team losses (red team kills).

Chi-square tests were performed on the blue team kills ($\chi^2 = 5.973$, df = 4) and blue team losses ($\chi^2 = .05$, df = 2) between trials in which there were stops to collect SAGAT data and those in which there were no stops, depicted in Figures 3 and 4. There was no significant difference in pilot performance ($\alpha = .05$) on either measure.

Analysis of variance was used to evaluate the effect of number of stops and duration of stops on each of the two performance measures -- blue team kills and blue team losses.

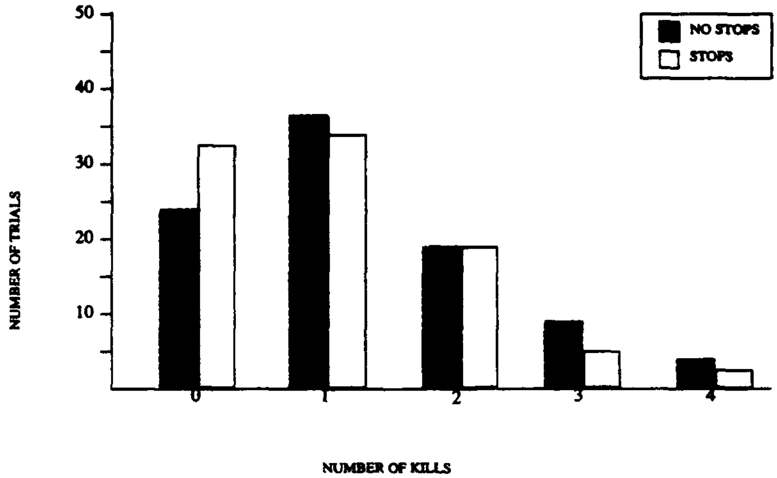


Figure 3. Blue Kills: Trials which include stops versus trials which were not stopped

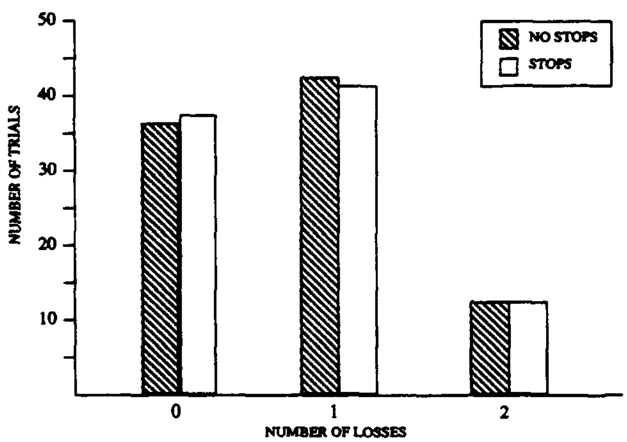


Figure 4. Blue Losses: Trials which include stops versus trials which were not stopped

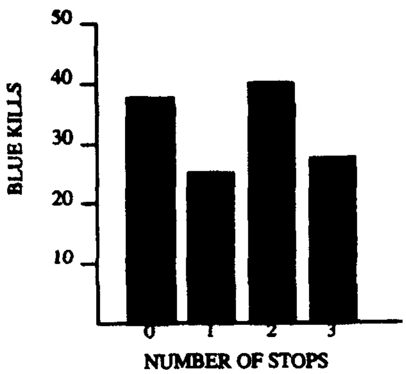


Figure 5. Blue kills by number of stops in trial

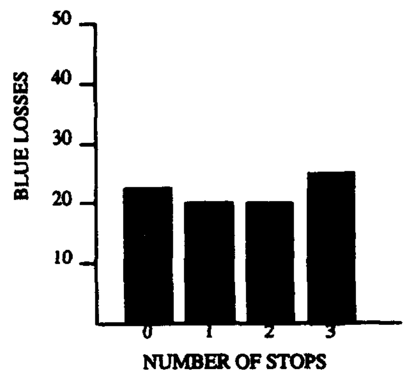


Figure 6. Blue losses by number of stops in trial

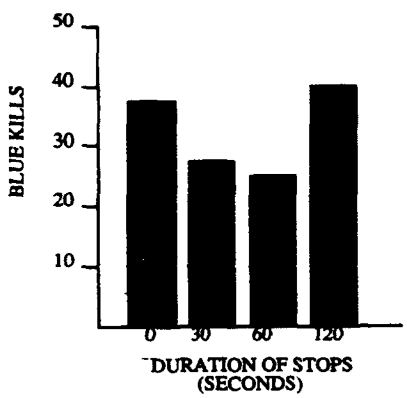


Figure 7. Blue kills by duration of stops in trial

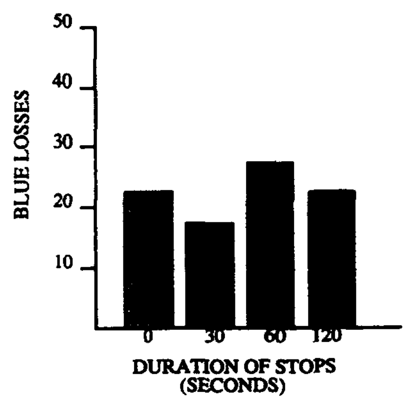


Figure 8. Blue losses by duration of stops in trial

The number of stops during the trial had no significant impact on either pilot performance measure at the $\alpha = .05$ level ($F=1.73$, $df = 3,116$ and $F=.20$, $df = 3,116$ respectively), shown in Figures 5 and 6. The duration of the stop also did not significantly impact either performance measure at the $\alpha = .05$ level ($F=2.16$, $df = 3,116$ and $F=.77$, $df = 3,116$ respectively), depicted in Figures 7 and 8. In viewing the data, no linear trend even appears to be present in either case that would indicate a progressively worse (albeit nonsignificant) effect of increasing number or duration of stops. This data supports the null hypotheses, indicating that stops to collect SAGAT data (as many as three for up to two minutes in duration) will not have a significant impact on later subject performance.

Discussion

The lack of a significant influence of this procedure on performance probably rests on the fact that relevant schema are actively utilized by subjects during the entire freeze period. Under these conditions, the subject's SA does not have a chance to decay before the simulation is resumed (as was indicated by Study 1). Thus, their SA is fairly intact when the simulation continues, allowing them to proceed with their tasks where they left off.

These results are being viewed with some caution, however. More such tests are probably needed to assess with certainty that the freeze and restart does not influence subsequent performance. Subjectively, the subjects did fairly well with this procedure and were able to readily pick up the simulation at the point where they left off at the time of the freeze, sometimes with the same sentence they had started before the stop. On many occasions, subjects could not even remember if they had been stopped to collect SA data during the trial, also indicating a certain lack of intrusiveness.

The recommendation for SAGAT administration at this point is that if SAGAT data is collected in this manner, some trials should be conducted during which SAGAT data is not collected, so that a check is provided for any influences that a freeze and restart in the simulation may cause. SAGAT may also be administered without restarting the trial afterwards, if assessment of SA is the primary objective and there is no desire for overall performance measurement.

Conclusions And Recommendations

In conclusion, the use of a temporary freeze in the simulation to collect SA data is supported by these two studies. Subjects were able to report their SA using this technique for as long as five or six minutes without apparent memory decay and the freeze did not appear to be intrusive on subject performance in the simulation, allaying several concerns about the technique. While it is always difficult to establish no effect of one variable on another (i.e., prove the null hypothesis), it is reassuring that the finding of no intrusion on performance has been repeated in numerous other studies where SAGAT was used to evaluate HMI concepts (Bolstad and Endsley, 1990; Endsley, 1989b; Northrop, 1988), as long as the stop is unpredictable to the subjects. This finding was even repeated in one study in which pilots flew in a simulation against computer controlled enemy aircraft (Endsley, 1990c). This helps to rule out the possibility that the finding of no impact on performance in the present study could have been the result of pilots on both sides of the simulation being equally impacted.

Several methods for measuring SA have been discussed, each with advantages and disadvantages that must be weighed carefully when selecting a metric. Ultimately, validity and reliability must be established for any SA measurement technique that is used. This, however, proves to be a very difficult task when there is no previously existing objective measure of the construct, other than the measure to be validated itself. This being the case, it is necessary to establish that the metric (a) indeed measures the construct it claims to measure and is not a reflection of other processes, (b) provides the required insight in the form of sensitivity and diagnosticity, and (c) does not substantially alter the construct in the process,

providing biased data and altered behavior. In addition, it can be useful to establish the existence of a relationship between the measure and other constructs as would be predicted by theory. In this case, one would probably need to establish that the measure of SA was predictive of performance and was sensitive to manipulations in workload and attention.

The SAGAT technique has thus far proven to meet these criteria. In addition to the present studies establishing a level of empirical validity, SAGAT has been shown to have predictive validity (Endsley, 1990a) and content validity (Endsley, 1990c). The method is not without some costs, however, as a detailed analysis of SA requirements is required in advance in order to develop the battery of queries to be administered. On the positive side, this analysis can also be extremely useful for guiding design efforts, going far beyond traditional task or information requirements analyses.

In addition to establishing measure validity, more research is needed regarding the task of SA measurement itself. For instance, as pointed out by Pew (1991), no criteria exist at this time establishing the level of SA which is required for successful performance. Does an operator need to have SA which is 100 percent perfect (in both completeness and accuracy), or is some lesser amount sufficient for good performance? This is a complex issue. I would assert that SA can be seen as a factor which increases the probability of good performance, but which does not necessarily guarantee it, other factors also coming into effect (decision making, workload, performance execution, system capabilities, SA of others in some cases). How much SA one needs therefore becomes a matter of how much probability of error one is willing to accept. Perhaps such criteria should more properly, and usefully, be established as the level of SA needed on each subcomponent at different periods in time. Some guidelines will eventually need to be specified if there is a desire to certify new system designs on the basis of SA.

Overall, the beginnings of the field of SA measurement have been established, allowing researchers in the area of SA to proceed from mainly speculation and anecdotal information to solid empirical findings. The tools for conducting further basic research exploring the SA construct and developing better system designs are available, allowing needed research on situation awareness in a variety of arenas.

Acknowledgments

This research was sponsored by the Northrop Corporation. I would like to thank the many people at Northrop, both past and present, who contributed to this effort, in particular, Vic Vizcarra for his assistance in conducting the studies, Paul Cayley and Gerry Armstrong for their programming efforts, and Mike Major and Del Jacobs for their support in developing a situation awareness research program. In addition, I am indebted to Mark Chignell who provided helpful support throughout the development of the methodology.

References

Air Force Studies Board Committee on Automation in Combat Aircraft. (1982). *Automation in combat aircraft*. Washington D.C.: National Academy Press.

AMRAAM OUE Final Report. (1982). St. Louis, MO: McDonnell Douglas Aircraft Corporation.

Bolstad, C. A., & Endsley, M. R. (1990). *Single versus dual scale range display investigation* (NOR DOC 90-90). Hawthorne, CA: Northrop Corporation.

Chambers, A. B., & Nagel, D. C. (1985). Pilots of the future: Human or computer? *Computer, No.*, 74-87.

Chase, W. G., & Simon, H. A. (1973). Perceptions in chess. *Cognitive Psychology, 4,* 55-81.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104*(2), 163-191.

Dornheim, M. A. (1986). Crew situation awareness drives avionics developments. *Aviation Week and Space Technology, June 23,* 114-116.

Endsley, M. R. (1987). *SAGAT: A methodology for the measurement of situation awareness* (NOR DOC 87-83). Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (1988a). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97-101). Santa Monica, CA: Human Factors Society.

Endsley, M. (1988b). Situation Awareness Global Assessment Technique (SAGAT). *Proceedings of the National Aerospace and Electronics Conference (NAECON)* . New York: IEEE.

Endsley, M. R. (1989a). *Final report: Situation awareness in an advanced strategic mission* (NOR DOC 89-32). Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (1989b). *Tactical simulation 3 test report: Addendum 1 situation awareness evaluations* (81203033R). Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (1990a). Predictive utility of an objective measure of situation awareness. *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 41-45). Santa Monica, CA: Human Factors Society.

Endsley, M. R. (1990b). *Situation awareness global assessment technique (SAGAT): Air-to-air tactical version user guide* (NOR DOC 89-58, rev A). Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (1990c). *Situation awareness in dynamic human decision making: Theory and measurement* (NOR DOC 90-49). Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (in preparation). *Situation awareness and workload: Flip sides of the same coin.* Manuscript in preparation.

Endsley, M. R., & Bolstad, C. A. (in preparation). *Individual differences in pilot situation awareness.* Manuscript in preparation.

Fracker, M. L. (1989). Attention gradients in situation awareness. In *Situation Awareness in Aerospace Operations* (pp. 6/1-1/10). Copenhagen, Denmark: NATO - AGARD.

Herrmann, D. J. (1984). Questionnaires about memory. In J. E. Harris & P. E. Morris (Eds.), *Everyday memory, action and absent-mindedness* (pp. 133-151). London: Academic Press.

Hughes, E. R., Hassoun, J. A., & Ward, F. (1990). *Advanced tactical fighter (ATF) simulation support: Final report* (CSEF-TR-ATF-90-91). Wright-Patterson, AFB, OH: Aeronautical Systems Division, Air Force Systems Command.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306-340.

Jurgensen, J. R., & Feldman, R. E. (1985). Expert system pilot aid. *Proceedings of the National Aerospace Electronics Convention (NAECON)* (pp. 1583-1590). New York: IEEE.

Kellog, R. T. (1980). Is conscious attention necessary for long-term storage? *Journal of Experimental Psychology: Human Learning and Memory, 6*, 379-390.

Manktelow, K., & Jones, J. (1987). Principles from the psychology of thinking and mental models. In M. M. Gardiner and B. Christie (Eds.), *Applying cognitive psychology to user-interface design* (pp. 83-117). Chichester: Wiley and Sons.

Marshak, W. P., Kuperman, G., Ramsey, E. G., & Wilson, D. (1987). Situational awareness in map displays. *Proceedings of the Human Factors Society 31st Annual Meeting* (pp. 533-535). Santa Monica, CA: Human Factors Society.

Morishige, R. I., & Retelle, J. (1985). Air combat and artificial intelligence. *Air Force Magazine, October*, 91-93.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76*, 165-178.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84* (3), 231-259.

Norman, D. A. (1968). Towards a theory of memory and attention. *Psychological Review, 75*(6), 522-536.

Northrop Corporation. (1988). *Tactical simulation 2 test report: Addendum 1 situation awareness test results.* Hawthorne, CA: Author.

O'Shannon, D. (1986). Expert systems in EW. *DS & E, March*, 19-25.

Person, L. H., & Steinmetz, G. G. (1981). *The integration of control and display concepts for improved pilot situation awareness.* Paper presented at the Flight Safety Foundation, 34th International Air Safety Seminar, Acapulco, Mexico.

Pew, R. W. (1991). Defining and measuring situation awareness in the commercial aircraft cockpit. *Proceedings of the Conference on Challenges in Aviation Human Factors: The National Plan* (pp. 30-31). Washington, D.C.: American Institute of Aeronautics and Astronautics.

Press, M. (1986). *Situation awareness: Let's get serious about the clue-bird.* Unpublished manuscript.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology, 1*(1), 45-57.

Selcon, S. J., & Taylor, R. M. (1989). Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design. In *Situation Awareness in Aerospace Operations* (pp. 5/1 -5/8). Copenhagen, Denmark: NATO - AGARD.

Stiles, G. J., & Pearson, R. P. (1986). Designing for situational awareness. *Defense Electronics, Dec*, 85-88.

Taylor, R. M. (1989). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situation Awareness in Aerospace Operations* (pp. 3/1 - 3/17). Copenhagen, Denmark: NATO - AGARD.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40*, 385-398.

Venturino, M., Hamilton, W. L., & Dvorchak, S. R. (1989). Performance-based measures of merit for tactical situation awareness. In *Situation Awareness in Aerospace Operations* (pp. 4/1 - 4/5). Copenhagen, Denmark: NATO - AGARD.

Vidulich, M. A. (1989). The use of judgment matrices in subjective workload assessment: The subjective workload dominance (SWORD) technique. *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 1406-1410). Santa Monica, CA: Human Factors Society.

Flightdeck

- **Tactical Cockpits – The Coming Revolution**
- **"How in the world did I ever get into that mode?" Mode Error and Awareness in Supervisory Control**
- **The Ubiquitous Three in the Prediction of Situational Awareness: *Round Up the Usual Suspects***
- **Comparison of Pilots' Acceptance and Spatial Awareness When Using EFIS vs. Pictorial Display Formats for Complex, Curved Landing Approaches**

Tactical Cockpits – The Coming Revolution

Eugene C. Adam

McDonnell Aircraft Company

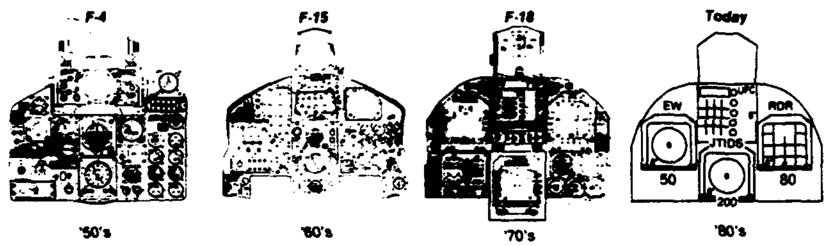
Introduction

Never has the cockpit designer had such a rich selection of emerging technologies from which to choose. But in these austere times, this treasure trove of technologies is under severe pressure to pay its way in combat kills, safety, or survivability. Therefore, each technology needs to be evaluated on the basis of which problem it solves and the cost effectiveness of the solution.

Before examining these new technologies it might be useful to first examine today's cockpits to see where we stand.

As shown in Figure 1, the analog cockpit of the two-place F-4 Phantom was followed by the HUD/CRT/Analog cockpit of the one-place F-15 Eagle which gave way to the HUD/multifunction display (glass) cockpit of the dual mission, one-place F/A-18 Hornet. Most of the western fighters built since that time use similar cockpit schemes: 1) a Head-Up Display, 2) Some Multi-Function Displays, 3) An Up-Front Control and 4) Hands on Throttle and Stick (HOTAS).

Cockpits have progressed from "steam gauges" to multipurpose displays.



However: *The greatest challenge facing today's cockpit designer* is to provide the pilot with the necessary Situation Awareness (SA) to be effective in combat. Todays cockpits have difficulty providing that SA because:

- Over 70% of Panel is inflexible
- Only 10 – 20% of Panel provides combat information
- Displays are too small to overlay Radar/NAV/EW/JTIDS on a map
- Display technology is stagnant because of low funding
- Pilot has no Head-Out Information except in the area of the HUD

GP03-0573-2-R

Figure 1. From round dials to multifunction displays. Where do we go from here?

Situational Awareness in Complex Systems
Edited by Richard D. Gilson, et al.

Two Problems

There are two major problems with applying today's cockpit to tomorrow's sensor/mission environment: 1) today's pilot spends more time managing and integrating sensors than executing Tactics and, 2) Useful combat information is available on only 20% - 30% of the instrument panel.

Fiddling and Flying

The first problem requires the pilot to fiddle around with a host of sensors and try to mentally integrate the data from the three primary ones while flying. Radar, EW, and data link are presently displayed on three different displays, on three different range scales with two or three different "ownship" locations. In the past, this has not been an overriding problem because radar search volumes were small and they generally tracked only a few targets, EW systems were inaccurate and full of false alarms and thus largely ignored, and JTIDS/Data Links were aboard very few aircraft. However, this will not be the case in the 21st Century. Sensor search volumes will increase at least one order-of-magnitude, EW accuracies will improve and data links will be common. These factors will greatly impact the pilot's ability to remain the "sensor manager/integrator" and have time left over for tactics execution.

Unproductive Space

The second problem, that of inefficient use of the instrument panel space, is a straight geometry equation. The average instrument panel is roughly 18" high by 24" wide or about 400 square inches. Using (3) 5" or 6" CRT's yield a total display area of 75 to 108 square inches. Therefore, on average, 70 to 80% of the instrument panel is inflexible, devoid of combat data, and unable to contribute to the fight or bombing run.

Since hostile contact generally averages only 30 seconds to 2 minutes, the pilot has to cope with unfused data on small displays on only a fraction of the instrument panel in a time-critical, high-stress, high-g environment. Not a good formula for making "everybody an Ace".

In combat, the pilot is in the aircraft to make good tactical decisions and execute them. Everything else is secondary. However, the correctness of tactical decision-making is directly proportional to the Situation Awareness (SA) of the pilot.

Situation Awareness (SA)

So, what is SA? What is it all about? It's simply KNOWING WHAT'S GOING ON SO YOU CAN FIGURE OUT WHAT TO DO! Where are the friendlies, bogies, SAM's and unknowns with respect to my flight? What are their intentions, my intentions and my options? It's obvious that present cockpits, by separating primary sensor data, on different

range scales with different "ownship" positions, do not give the pilot the SA required to achieve the exchange ratios necessary to win against superior numbers of equivalent quality targets.

The Big Picture

As shown in Figure 2, SA is a two-fold problem: Global and Tactical. Global SA (the Big Picture) generally covers the non-visual spherical world at ranges from 0 to 200 miles. Most often a plan view SA is best, with your ownship position decentered because of higher interest and lethality in the forward hemisphere. However, even in a low-intensity conflict, the 100 mile range display could contain hundreds of graphic elements such as unfriendly surface and airborne threats, friendly surface and airborne elements, unknowns, navigation paths, map and symbolic data. Separate, small displays are no match for this complexity.

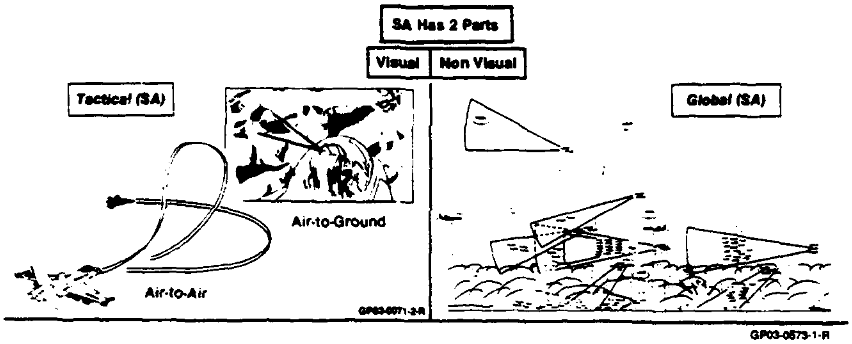


Figure 2. Situational Awareness – "Knowing what's going on so you can figure out what to do."

The Little Picture

Tactical SA covers close-in visual air-to-air and air-to-surface combat and visual navigation. M on N combat is the arena where man and machine are taxed to their limits. For equivalent machines, the SA acted upon by the eye, brain, hands and feet is the primary determinant of "who shoots" and "who chutes".

Tactical SA Solution

The tactical SA problem is best solved by a helmet system that: 1) TRACKS the pilot's head position and slaves sensors and weapons to the helmet line-of-sight, 2) DISPLAYS combat and flight information on the helmet visor.

Integrated Helmet System. MCAIR and Kaiser Electronics IRADs have designed, built, simulator tested and flown an Integrated Helmet Mounted Display and Sight (HMDS) System called "Agile Eye" (TM) which can increase visual exchange ratios by a factor of 2:1 over a Head-Up Display. The "Agile Eye" is a totally integrated helmet sight and display that has the following features:

- A HUD type display on the visor
- Lighter than present helmets
- Improved CG
- Improved crash protection,
- No visual obstructions,
- Less aerodynamic lift during ejection,
- Improved sound reproduction/attenuation.

The "Agile Eye" Helmet uses readily available off-shelf technology cleverly integrated into a pilot centered design that improves every physical and performance characteristic of today's flight helmet. It offers fields-of-view and stroke/raster capabilities that match present day HUD's but with the advantages of off-axis weapon use, three quarters of the system cost, two times the reliability and the added safety of attitude and other flight data available at all times, and at all sight angles. All of these features are packaged in a low-bulk, handsome design as pictured in Figure 3.

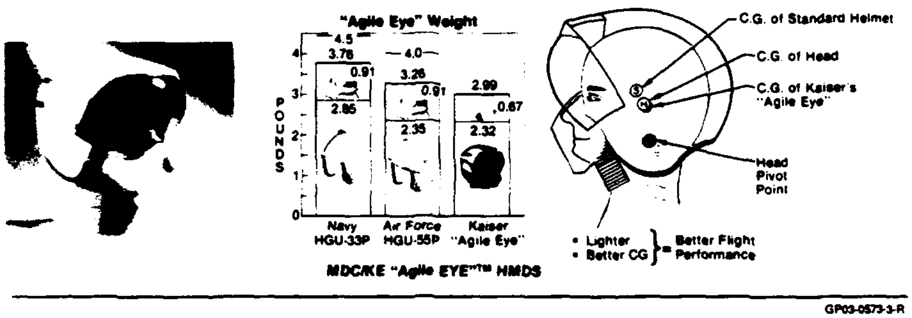


Figure 3. "Agile Eye" – a HUD-on-the-head without penalties to the pilot.

"Agile Eye" Payoff. In A/A, there are faster visual lock-ons, simultaneous AIM-7 and AIM-9 launches, target hand-offs to wingman, better attitude awareness at all times. In A/G, there are off-boresight target designations, offset NAV waypoint updates, target hand-offs to

wingman. As shown in Figure 4, MCAIR F-15 simulator evaluations using TAC pilots/aggressors/scenarios showed a 2:1 exchange ratio improvement with the "Agile Eye" HMDS over the HUD.

Helmet Systems - The Linchpin. We are convinced that helmet systems are the key to future cockpit improvements; they increase a pilot's performance and free-up panel space.

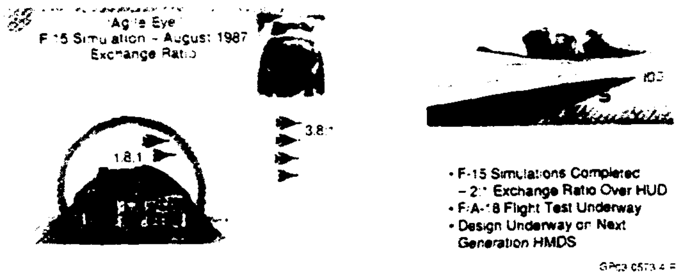


Figure 4. "Agile Eye" Doubtless Exchange Ratios

What is the Pilot's Problem?

The pilot's problem is threefold: 1) Next generation sensors, such as RADAR, EW and JTIDS will provide 100's of pieces of information; 2) current display technology limits CRT sizes to 5, 6 or 7 inches square; and 3) small displays require separation of primary sensors such as RADAR, EW, JTIDS and NAV, leaving the pilot to mentally integrate and fuse this data during the stress of combat.

Picture This!

Three different sensors on three separate displays on three different range scales with "ownship" in three different locations equals a formula for confusion. Larger displays solve that problem by fusing all sensor data to a common range scale and coordinate system and overlaying it on a map.

What is the Hardware Problem?

CRTs using a scanning beam naturally grow dimmer as they are made larger, a fact which is unacceptable in a high ambient cockpit. Flat panel displays using matrix addressed pixels

do not have this problem, but the technology and infrastructure need R&D funds before they can seriously challenge the CRT.

Global SA Solution

The beyond-visual-range Situational Awareness solution requires the "fusion" of RADAR, EW, JTIDS navigation and map on a large display. This would allow the pilot to look at a single source to "get the Big Picture".

As shown in Figure 5, display size growth has not kept pace with computer and sensor technology because of the lack of serious research and development on CRT alternatives. A two-step solution offers the most cost and schedule effectiveness. In the near term, we must first develop larger, new technology displays on which to display the situation to the pilot. We must then reconfigure the HUD to provide the room to mount this display in the cockpit. In the far term we must develop new, flat-panel matrix technologies that provide display surfaces of 10 to 15 times what is available using today's CRT technology.

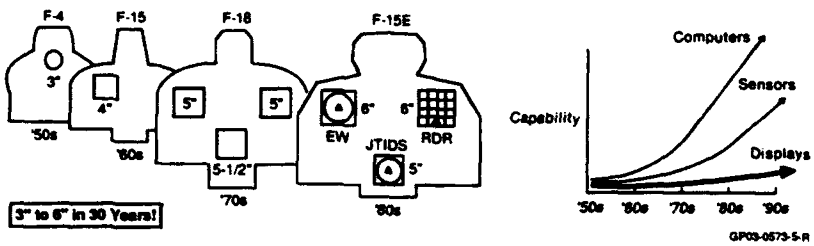


Figure 5. Present evolution of displays not keeping up with computers and sensors.

Cockpit 2000: A Near Term Solution

Helmet systems such as "Agile Eye" are essentially a HUD-on-the-head which allows us to reduce the physical size of the aircraft HUD sufficiently to provide room for a 10" x 10" Global Situation Display. This display is a compromise between being large enough to fuse RADAR, EW and JTIDS on a single touch sensitive surface, but yet small enough to leave room for adjacent 5" or 6" auxiliary displays.

As shown in Figure 6, Cockpit 2000 has about 2X the display area of current fighters and differs from today's cockpit in two important aspects: 1) A helmet sight and display provides all normal HUD functions on the helmet visor with the added benefit of off-axis target designation, and 2) The 10" x 10" Global Situation Display is larger and more productive than any three, small multifunction displays.

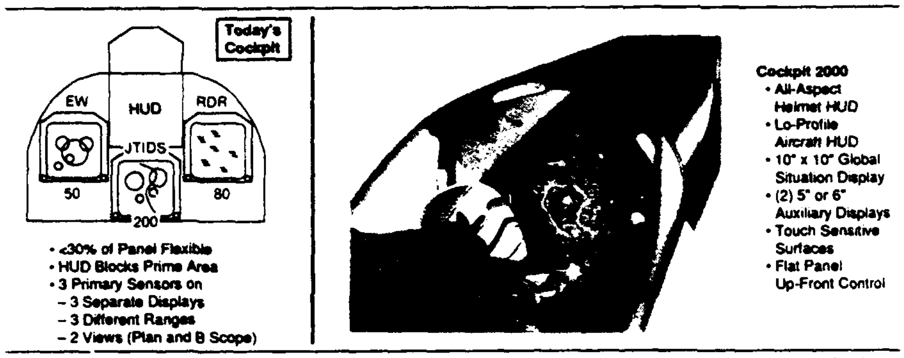


Figure 6. Cockpit 2000 solves the two most pressing cockpit problems: Tactical and Global SA.

Big Picture: A Longer Term Solution

An increase in display technology R&D will eventually provide flat, matrix display panels with large surface areas, high brightness, high resolution and long life. As depicted in Figure 7, these large displays will provide 10 times the display area of today's CRTs allowing plan and perspective views, split screen, and movable inserts. A Helmet Sight and Display, voice command and touch sensitive surface will provide pilot interface with the weapon system. In short, the Big Picture provides the pilot with full control over the configuration and content of almost 400 square inches of display surface to match the mission-moment-of-interest whether it be *air-to-air*, *air-to-surface*, Navigation, TF/TA, or System Status. Manned Simulations have shown a 100% increase in the situational awareness of pilots using the Big Picture over those using a conventional 2 or 3 small MFD (CRT) cockpit.

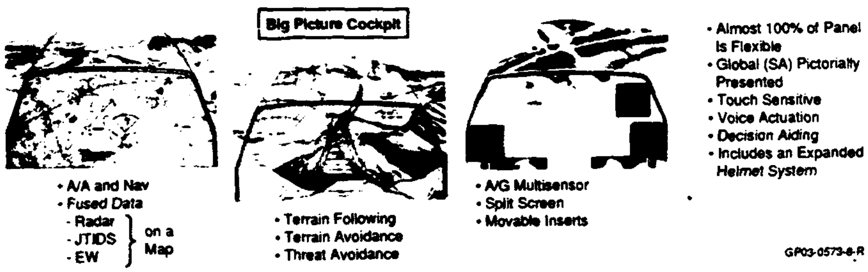


Figure 7. "Big Picture" provides total reconfiguration flexibility.

Display Technology

The CRT has reigned supreme as the display device of choice for almost 100 years, with continuous evolutionary brightness, resolution, reliability and color improvements over that time. In fact, the huge CRT infrastructure and its good performance has stifled any real competitive technology investments until recently.

There are three large markets for a CRT replacement: 1) HDTV promises displays sizes of 2-5 times present CRT devices with the desire to "hang it on the wall" like a picture. 2) Portable PC's up through work stations desire high-resolution, full color, small bulk and for portable applications, low-power consumption. 3) Military and Aerospace all share a similar problem; too much data on too small a CRT surface. Larger displays are required to solve this problem but the bright sunlight conditions in aircraft must also be met which essentially dooms the CRT.

All three of these applications and their commercial profit potential are giving a massive push to flat panel technologies. The next three years will see a R&D investment in flat panels of at least three times the total CRT alternative investments for the last 30 years. Unfortunately, the U.S. investment is roughly 5% of the worldwide investment, hence our commercial possibilities are few and our defense needs may well be supplied by offshore manufacturing facilities.

The Final Frontier

The laser, CBR and high energy weapon threat will require radical approaches to protecting the crew and providing sufficient information to fly and fight. There are two broad alternative solutions: 1) Remove the crew from the cockpit and fly and fight using remotely controlled vehicles, and 2) protect the crew within a "windowless cockpit".

Remotely Piloted Vehicles

Two technologies are necessary to provide this capability: 1) Sensors equivalent to the eye/brain are required to capture the visual combat scene real-time. 2) A secure, wide-bandwidth data link is required with near real-time capability to allow a pilot to fly and fight from a remote location.

For convenience, we will not treat this case because SAM's, cruise missiles and other weapons fill many of these mission functions and the technology and frequency spectrum required for the immense amount of data to be linked between the pilot and vehicle on a real-time non line-of-sight basis make it impractical for any large number of fighters.

Windowless Cockpit

Needless to say, the concept of a sleek fighter without a canopy will cause most pilots to shudder and gag. However, the laser threat is real, they are in the field and 50 mile, zero

time-of-flight "dazzles" are on the horizon. For simplicity, let us assume that sensors can provide spherical coverage around the aircraft with visual acuity. With the windowless cockpit concept there are two broad solutions: 1) Retractable protection whereby the pilot flies visual or non-visual depending on the situation and trains both ways; and 2) full time, enclosed cockpits with no outside vision. Both solutions require helmet displays and fixed displays; however, the retractable protection scheme has the disadvantage of having to meet 1000 times the ambient brightness requirement of the fully enclosed alternative.

Helmet vs. Cockpit Displays

Without enormous breakthroughs in optics and display devices, the goal of a helmet display that does everything and doesn't require additional head-down displays does not seem practical for the high g environment in the near term. As shown in Figure 8, Cockpit and Helmet Displays are complementary. Both are required and both need extensive R and D to meet the needs of all three generations of cockpits discussed herein.

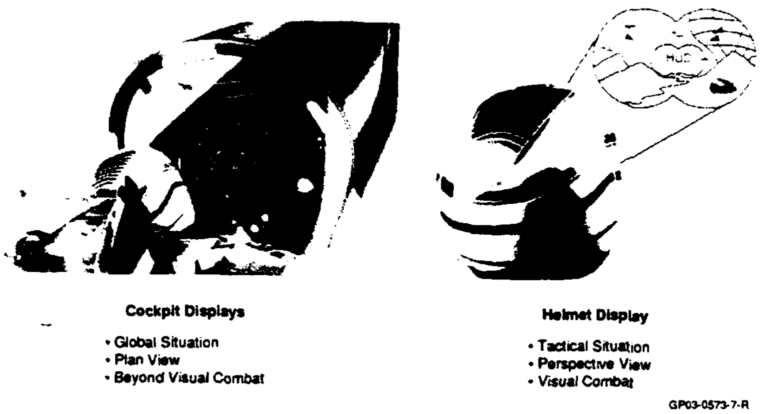


Figure 8. Cockpit and helmet displays, they complement each other. Both are required for SA.

Supporting Technologies

A number of supporting technologies are needed to gain the full advantages of the three generations of cockpits proposed herein. The real issue, however, is the cost/benefit ratio of individual and combined technologies. These are difficult questions to answer definitively because simulations and tests tend to emphasize environments whereby tested technologies are useful when nobody knows what the eventual distribution of scenarios will actually be. Fortunately, the aerospace industry and DOD. have seasoned design teams that are very good at getting the right systems in the final version of new generation aircraft.

Summary

The HUD and Multifunction Display cockpit using 5" and 6" CRT's have served us well. They have, however, two weaknesses: 1) No off-axis designation and information but this can be solved with Integrated Helmet Systems, and 2) No fused sensor and NAV data to a common range and coordinate system. This solution requires a large display, which most likely will be a non CRT technology.

The 90's will see a juncture of technologies such as flat panels, speech, graphics, decision aids, and immense computational capability ripening for the cockpit designers picking. Mission and vehicle requirements will and should drive the final choices.

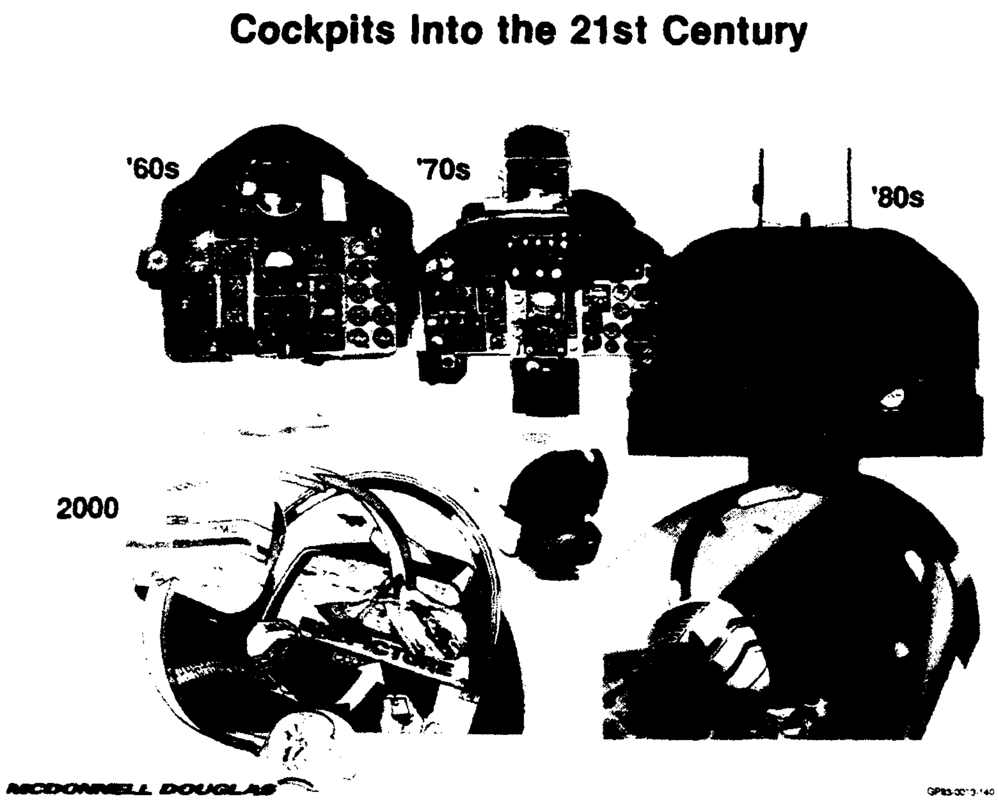


Figure 9. The evolution of cockpits in the 20th century.

"How in the world did I ever get into that mode?" Mode Error and Awareness in Supervisory Control

Nadine B. Sarter & David D. Woods

The Ohio State University

Introduction

New technology is increasing the potential for automated resources to support human supervisory controllers. The technology's inherent flexibility allows designers to add a wide range of capabilities in the name of providing the practitioner with a set of tools that can be used to optimize system performance across a wide range of circumstances. However, the same flexibility tends to create and proliferate various modes of operation. This proliferation of modes that can so easily accompany new levels of automation in complex systems also creates new cognitive demands on practitioners (Woods, 1993). Practitioners must know more – both, about how the system works in each different mode and about how to manage the new set of options in different operational contexts (Sarter and Woods, 1992; 1993). New attentional demands are created as the practitioner must keep track of which mode the device is in, in order to select the correct inputs when communicating with the automation and in order to track what the automation is doing now, why it is doing it, and what it will do next. For example, an automated cockpit system such as the Flight Management System (FMS) is flexible in the sense that it provides pilots with a large number of functions and options for carrying out a given flight task under different circumstances.

There are at least five different methods at different levels of automation that the pilots could invoke to change altitude. This flexibility is usually portrayed as a benefit that allows the pilot to select the mode best suited to a particular flight situation. But this flexibility also has a price: the pilots must know about the functions of the different modes, how to coordinate which mode to use when, how to 'bumplessly' switch from one mode to another, how each mode is set up to fly the aircraft, and to keep track of which mode is active. These new cognitive demands can easily congregate at high tempo and high criticality periods of device use, thereby adding new workload at precisely those time periods where practitioners are most in need of effective support systems. Clumsy use of technological possibilities, such as the proliferation of modes, creates the potential for new forms of human-machine system failure and new paths towards critical incidents (e.g., the air crashes at Bangalore) (e.g., Lenorovitz, 1990) and Strasbourg (Monnier, 1992).

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

In a variety of studies, we have investigated human-automation interaction in the context of commercial 'glass' cockpits (Sarter and Woods, 1992; 1993) and in the context of anesthetic management under surgery (Cook et al., 1990; Moll van Charante et al., 1992). Based on these and other studies, we think that the classic concept of mode error is inadequate to describe the problems in human interaction with today's automated resources. In this paper, we extend the concept of mode error to take into account the problems caused by new automation capabilities – mode awareness.

Mode Awareness

Multiple modes in devices can create the potential for mode errors. The concept of mode error has been established as one kind of problem that can occur in human interaction with computerized devices (Lewis and Norman, 1986) and as a basic kind of erroneous action in psychological taxonomies of error forms (Norman, 1981). Norman (1988) summarizes the source of mode error quite simply by suggesting that if one wishes to create or increase the possibilities for erroneous actions, one way is to "... change the rules. Let something be done one way in one mode and another way in another mode." When this is the case, a human user can commit an erroneous action by executing an intention in the way appropriate to one mode of the device when the device is actually in another mode.

Note that mode error is inherently a human-machine system breakdown in that it requires that the users lose track of which mode the device is in (or confuse which methods or actions are appropriate to which mode) and it requires a machine where the same actions and indications mean different things in different modes of operation. Several studies have shown that human-computer interface design and evaluation should identify computerized devices which have a high potential for mode errors (e.g., Lewis and Norman, 1986; Cook et al., 1991), and several design techniques have been proposed to reduced the chances for mode errors (Monk, 1986; Sellen et al., 1992).

The original work on mode error was done primarily in reference to relatively simple computerized devices, such as word processors. The erroneous actions in question were acts of commission in carrying out self paced tasks with devices that only reacted to user inputs and commands. Increases in the complexity and autonomy of automated systems for event-driven, dynamic task environments, such as commercial aviation flightdecks and anesthetic management under surgery, have resulted in a proliferation of system and interface "modes." Human supervisory control of automated resources in event-driven task domains is a quite different type of task environment as compared to the applications in the original research on mode error. Automation is often introduced as a resource for the human supervisor providing him with a large number of modes of operation for carrying out tasks under different circumstances. The human's role is to select the mode best suited to a particular situation, but to accomplish this he or she must know more and must meet new monitoring and attentional demands to track which mode the automation is in and what it is doing to manage the underlying process. These cognitive demands can be particularly challenging in the context of highly automated systems which can change modes on their own based on environmental inputs or for protection purposes, independent of direct and immediate instructions from the human supervisor. This capability of highly automated systems drives the demand for mode

awareness, that is, the ability of a supervisor to track and to anticipate the behavior of automated systems.

What is involved in maintaining mode awareness is determined to a large extent by the design and capabilities of the automated resources and especially the interface between the automation and the people in the system. Therefore, how have changes in automation and in the interface between person and automated resources impacted mode awareness? How has the human's role and tasks changed and how can they be supported?

The Complexity of Modes in Automated Systems and the Challenge to Mode Awareness

Early automated systems were characterized by a fairly small number of modes. In most cases, these modes provided the passive background on which the operator would act by entering target data and by requesting system operations. Another characteristic of these early systems was that they would only have one overall mode setting for each function to be performed. Consequently, mode annunciations (indications of the currently active mode and of transitions between modes) could be dedicated to one spatial location on the display. Finally, consequences of a breakdown in mode awareness tended to be fairly small, in part because of the short time-constant feedback loops involved in these systems. The operator seemed to be able to detect and recover from erroneous actions relatively quickly.

The flexibility of more advanced technology allows automation designers to develop much more complicated mode-rich systems. Modes proliferated by providing multiple levels of automation and by providing more than one mode option for many individual functions. The result is numerous mode indications distributed over multiple displays each containing just that portion of the mode status data corresponding to a particular system or subsystem. Furthermore, the designs allow for interactions across the various modes. The increased capability of the automated resources themselves creates increased delays between user input and feedback about system behavior. This increase to longer time-constant feedback loops increases the difficulty of error or failure detection and recovery and challenges the human's ability to maintain awareness of the active modes, the armed modes, the contingent interactions between environmental status and mode behavior, and the contingent interactions across modes.

A very important trend relates to the input sources that can evoke changes in system status and behavior. Early systems would change their mode status and behavior only in response to operator input. Advanced technology, on the other hand, responds to operator input as well as situational and system factors. In the case of the Flight Management System in highly automated cockpits, for example, a mode transition can occur as an immediate consequence of operator input. But it can also happen when a preprogrammed intermediate target (e.g., a target altitude) is reached or when the system changes its mode in order to prevent the pilot from putting the aircraft into an unsafe configuration.

Even the aspect of operator input has itself become more complicated as the complexity of the system of automated resources has increased. Incidents and accidents have shown that there is an increased risk of inadvertent activation of modes by the operator. A mode can not only be activated through deliberate explicit selection of the mode by pushing the

corresponding button. In addition, pushing a button can result in the activation of other different modes depending on the system status at the time of manipulation. The resulting system behavior can be disastrous but may be missed by the operator if adequate feedback is not provided to support mode awareness.

An example of such an inadvertent mode activation contributed to a major recent accident in the aviation domain (the Bangalore crash; e.g., Lenorovitz, 1990). In that case, the pilot put the aircraft into a mode called OPEN DESCENT without realizing it. This resulted in the aircraft speed being controlled by pitch rather than thrust, i.e., throttles went to idle. In that mode, the automation ignores any preprogrammed altitude constraints. To maintain the pilot-selected target speed without power, the automation had to use an excessive rate of descent which ultimately led to the crash of the aircraft short of the runway. How could this happen?

There are at least three different ways of activating the OPEN DESCENT mode. First, it can be selected by pulling the ALTITUDE knob after selecting a lower altitude. Second, it can be activated by pulling the SPEED knob provided the aircraft is in the so-called EXPEDITE mode at that point in time. And third, the OPEN DESCENT becomes active when selecting a lower altitude while in the ALTITUDE ACQUISITION phase. This latter indirect option contributed to the above accident. The pilot must not have been aware of the fact that the aircraft was within 200 feet of the previously entered target altitude (which puts the system into the ALTITUDE ACQUISITION mode). Consequently, he may not have expected the selection of a lower altitude at that point in time to result in a mode transition. As he did not expect any mode change, he may not have closely monitored his mode annunciations and thus missed the transition. It was not until 10 seconds before impact that the crew discovered what had happened; too late for them to recover with the engines at idle.

Display modes are another factor aggravating the problem of mode awareness. In some devices, the current mode configuration does not only determine what control functions become activated by a given input; rather, these devices also interpret user-entered target values differently depending on the active display mode. In the following example, it is easy to see how this may result in unintended system behavior. In a current glass cockpit aircraft, pilots enter a desired vertical speed or a desired flight path angle via the same display. The interpretation of the entered value depends on the active display mode. But although the verbal expressions for different targets differ considerably (for example, a vertical speed of two thousand five hundred feet versus a flight path angle of two-point-five degrees), these two targets on the display look almost the same. The pilot has to verify the mode indication for this display instead of the display format supporting an intuitive, mentally economical apprehension of the active mode. In this case, the problem is further aggravated by the fact that the pilot is increasingly removed from the actual ongoing process as previously available cues about system behavior such as moving throttles or noise may have been reduced or removed in the design process.

The behavior and capabilities of the machine agent in human-machine systems have changed considerably. In simpler devices, each system activity was dependent upon operator input; consequently, in order for a lack of mode awareness to become operationally significant, the operator had to act to evoke undesired system behavior. In more automated systems, the level of animacy of machine agents has dramatically increased. Once activated, systems are capable of carrying out long sequences of tasks autonomously. For example, advanced Flight Management Systems can be programmed to automatically control the aircraft from takeoff through landing. Inadvertent mode settings and selections may not produce visible consequences for a long time complicating the process of error or failure

detection. This creates the possibility of errors of omission (i.e., failure to intervene) in addition to errors of commission as a consequence of a lack of mode awareness.

Another complicating factor that makes it difficult to maintain awareness of the active mode configuration is the fact that many systems allow for simultaneous use by multiple practitioners rather than input by just one individual user. Tracking system status and behavior becomes more difficult if it is possible for other users to interact with the system without the need for consent by all of the practitioners involved. This problem is most obvious when two experienced operators have developed different strategies of system use. When they have to cooperate, it can be particularly difficult for them to maintain awareness of the history of interaction with the system which may determine the effect of the next system input.

All of the factors mentioned above challenge a human supervisor's ability to maintain mode awareness in highly automated systems. The results of a number of studies of human-automation interaction in a variety of domains have indicated that problems in mode awareness are often a consequence of technology centered automation (e.g., Sarter and Woods, 1992 and 1993; Cook et al., 1991; Moll van Charante et al., 1992). In the following section, we will examine in more detail results of a series of studies on pilot-automation interaction that illustrates the trends in mode awareness problems in the context of the mode-rich cockpit environment.

Some Empirical Results on Mode Awareness in Pilot-Automation Interaction

The role of pilots in modern glass cockpit aircraft has shifted from direct control of the aircraft to supervisory control of automated machine agents. One of the core automation systems in these cockpits is the Flight Management System (FMS) which can be programmed by the pilot to automatically follow a desired flight path and profile from takeoff through landing. To maintain awareness of the status and behavior of the various modes of operation within the FMS, pilots have to gather and integrate a variety of data from numerous different displays in the cockpit. In addition to monitoring these nominal indications of system targets and status, pilots need to be able to interpret the indications to extract what is implied about current and future system and aircraft behavior. In other words, the automation is becoming more of a dynamic process in itself, where the indications are a kind of 'raw' data which require an act of interpretation in order to become information about current or future states. Interpreting the raw indications requires the human supervisor to have an adequate mental model of the various automated modes, their inter-relationships, and knowledge of how to use these as resources in various contexts.

Given the fairly low rate of change in aircraft behavior throughout large parts of the flight, the pilot does not have to continuously monitor the mode annunciations. Rather, he or she has to be able to predict the occurrence of transitions in system behavior to attend to the right indications at the right time. During busy phases of flight (e.g., final approach), numerous changes in system status and behavior can occur in a very short period of time. During this high tempo phase of flight, with a large number of concurrent tasks, the crew now has another

set of cognitive tasks to perform – monitoring and interpreting mode annunciations relative to expected behavior.

In a series of studies of pilot-automation interaction, we had the opportunity to investigate the nature and circumstances of mode-related problems in highly automated glass cockpit aircraft. In one investigation, we built a corpus of FMS-related problems that were encountered in line operations (Sarter and Woods, 1992). For this purpose, descriptions of automation surprises were collected from experienced airline pilots. A second converging activity was to observe pilots during their transition training from a conventional to a glass cockpit aircraft (i.e., before they had a chance to adapt to the system). Analysis of these corpus data suggested that difficulties in mode awareness and gaps in pilots' understanding of all of the modes of operation and their interactions contributed to automation surprises and related supervisory control difficulties. Based on the results from the corpus gathering studies, a field experiment was carried out to try to examine pilot-automation interaction more closely (Sarter and Woods, 1993). Twenty airline pilots were asked to fly a mission on a part-task flight simulator. The scenario was designed to contain numerous tasks and events that served as probes of pilots' mode awareness and of their mental model of the automation. This phenomenon-driven scenario design permitted on-line data collection on various issues regarding pilot-automation interaction. In addition, we were able to question the pilots about their knowledge and assessments of the status and behavior of the automation during low workload phases of the simulated flight and after completion of the simulation.

These studies provided consistent and converging data to help understand why and under what circumstances pilots encounter problems related to the interaction with cockpit automation. Most of the observed difficulties were related to lack of mode awareness and to gaps in mental models on how the various automated modes work and interact. The problems in coordination between pilot and automation (e.g., automation surprises) occurred primarily in the context of non-normal, time-critical situations; for example, aborted takeoff, disengaging an automatic mode during approach for collision avoidance, and loss of the glideslope signal during final approach. In the case of the aborted takeoff, 65% of the pilots were not aware that the autothrottle system was in charge of thrust control. Consequently, they did not disengage the autothrottles in order to have full manual control of the throttle setting. In the debriefing, 15% of these pilots could describe the active mode settings and the system activities during takeoff. But their knowledge was inert, i.e., they had not been capable of applying this knowledge to the ongoing situation. Overall, only four out of twenty participants responded completely correctly in managing the automation during the aborted takeoff, and one of these four pilots explained that he did so because he was trying to comply with standard procedures, not because he understood what was going on within automation. In the second non-normal situation, the pilots had to quickly comply with an ATC request to disengage the automatic APPROACH mode after localizer and glideslope capture in order to change heading and altitude to avoid a collision. Most of the pilots knew only one of the several methods to disengage the mode, and fourteen pilots also 'knew' at least one inappropriate method which could lead to delayed responses to the ATC request. In the case of the glide slope loss during final approach, about one half of the pilots were not aware of the consequences of this event in terms of FMS behavior. They could not explain the effects in the debriefing, and many of them even had difficulties detecting the occurrence of the problem during the ongoing simulation.

Another interesting result of this study was related to the future-oriented aspect of mode awareness. Pilots sometimes had problems anticipating system behavior and the associated mode annunciations. For example, only five of the participants knew when to expect the

indication that the Go-Around mode would be available. Failures to anticipate mode status and transitions, like this one, indicate a lack of mode awareness which degrades the pilot's ability to allocate attention effectively and to detect errors, failures, or miscommunications between pilot and automation prior to explicit flight events – automation surprises. The more experience pilots had with the automation, the more they were capable of applying their knowledge about the advantages and shortcomings of the different modes to manage the automated resources in different contexts. Pilots with less glass cockpit experience tended to utilize a single strategy or mode over a wide range of flight circumstances. One could interpret this as an attempt to cope with the complexity of the automation by ignoring some modes and options, even in situations where the stereotypical strategy was less than ideal relative to other strategies for managing the automated resources. Finally, there were several instances of pilots who instructed the automation by entering new flight path related targets but who did not activate a mode of the automation to work on acquiring these targets. They were surprised when the aircraft did not respond as expected; they did not realize or understand why their instructions to the automation had not resulted in the desired change. In some sense, this is a good example to show how pilots try to communicate with the system in a way analogous to communication with another human agent. They assume that entering a desired target value is sufficient for the system (as it would be for a human crewmember) to understand that it is supposed to achieve this new target and how it is supposed to do so in detail.

These investigations into one specific field of activity illustrate a trend in human-machine cooperation (Woods, 1993). Technology allows a proliferation of modes of increasing complexity and capability for autonomous activity. These changes create new cognitive demands for human supervisory controllers, demands which tend to congregate at higher tempo epochs where workload demands are highest (cf., also Moll van Charante et al., 1992 for similar results from another field of practice). The complexity of modes challenges the human supervisor's ability to track and anticipate the behavior of the automation – mode awareness. Difficulties in maintaining mode awareness focus on transitions between more quiescent phases or situations where mode behavior is complex or transitions frequent.

Sources of Problems in Mode Awareness

The data on problems in mode awareness imply that there are two kinds of contributing factors. One is buggy mental models. The other is opaque indications of the status and behavior of the automation. The former derives from a failure of the designers of automation to anticipate the new kinds of knowledge demands their automation creates and a failure to provide mechanisms to help practitioners acquire and maintain this knowledge in ways usable in actual operational contexts. The latter derives from a failure of designers to support the supervisor's increasingly challenging cognitive demand of tracking the state and behavior of the automation as another kind of dynamic process within their scope of responsibility (e.g., Norman, 1990). The indications of the nominal status of the automation are a kind of data; the practitioner must interpret this data to develop and maintain an assessment of the automation as process and the automated process over time. The data on problems in mode awareness strongly imply that this cognitive demand is poorly supported by the kinds of displays on the state of the automation currently provided to practitioners. As Earl Wiener likes to put it: the three most commonly asked questions on the highly automated flightdeck are – what is it doing? why is it doing that? what will it do next? (to which we would like to add a fourth –

how in the world did we get into that mode?). The interpretation of data on the automation as process is apparently a cognitively demanding task with these displays rather than a mentally economical one. This is troublesome when this cognitive task is important during high tempo, high workload, high criticality situations.

Coping with Mode Error and Aiding Mode Awareness

The examination of mode awareness here leads us to several strategic directions for responding to problems in this cognitive task. First, one can say that mode awareness problems are induced by the complexity of the technological system. Technological powers for automation are used clumsily when the cognitive and other kinds of demands on the operational system created by new automation are ignored (Norman, 1990; Woods, 1993; Woods et al., 1993). This is what we mean by technology-centered automation (cf., Billings, 1991 for an extensive discussion of technology-centered versus human-centered automation). Then one avenue to improve the human-machine system is to reduce the operational complexity induced by how technology is deployed. In the case of mode awareness this can be stated quite clearly – reduce the number and complexity of the modes. However, there may be a variety of pressures, such as marketing demands from a diverse set of customers or the methods for optimizing various parameters across different operational circumstances, which reduce the designer's ability to counter mode proliferation.

Two other directions for change probably are very tightly coupled in their implementation in a real field of practice, although they can be discussed separately. One is to support the new knowledge demands created by increasingly complex automated resources through new approaches to training human supervisory controllers. This is much more than simply a new list of facts about how the automation works. Instead, it must be focused on knowledge activation in context in order to avoid what we are already seeing – inert knowledge where the user can recite the facts but fails to use the knowledge effectively in real operational contexts. Training to enhance skill at control of attention would also be relevant here (Gopher, 1991).

The new knowledge demands require that more attention be paid to developing and teaching knowledge and strategies for how to work the system of automated resources in varying operational contexts. Finally, the knowledge demands of new levels of automation are strongly conditioned by a major constraint: if the automation is well engineered in a narrow sense, it will work well in a variety of routine situations; the problems of supervisory control will be manifest in situations with complicating factors that go beyond these routines. However, these will be relatively infrequent, at least for the individual practitioner. Thus, meeting the knowledge demands will require investing in maintaining usable knowledge relevant to the more difficult but infrequently occurring situations. In this, as in many other cases of introducing new levels of automation (e.g., Adler, 1986; Bereiter and Miller, 1988), new automation produces new training requirements.

A third direction for change is to develop new forms of aiding mode awareness itself through changes in the interface and displays that reveal what the automation is doing, why it is doing that, and what it will do next. The strategy is to provide better indications of what mode the system is in (to avoid mode errors), how future conditions may produce changes without direct practitioner intervention, and support better detection of and recovery from mode misassessments and mode errors when they do occur. Some attempts to do this have been made by changing the overall format of a display in different modes or changing the

cursor shape as the system transitions between modes. However, since the practitioner's visual channels are often heavily loaded in some fields of practice, signaling mode changes through non-visual channels such as aural or kinesthetic feedback may be useful (cf., Monk, 1986 and Sellen et al., 1992 respectively). Another concept is "history" displays of instructions to and of the behavior of automated systems. Such displays would provide a visual trace of past and projected system behavior under the current mode configuration. However, such displays would have to be 'intelligent' in that the future behavior of the automated systems is contingent on future events in the environment.

While there are several suggestions that can be offered on potentially fruitful directions ranging from general strategies on what are effective ways to clearly signal mode status and changes (e.g., use orienting perceptual channels such as auditory or tactile signals; Monk, 1986; Sellen et al., 1992) to particular tips (e.g., display active targets with mode annunciations), it turns out that the human error, cognitive engineering and human-computer interaction communities have barely begun to study the relevant issues to provide the necessary research base to drive or support practical advice to designers. However, developing such aids probably requires that we advance our understanding of how attention shifts across the perceptual field in dynamic multi-task domains (e.g., Eriksen and Murphy, 1987; Jonides and Yantis, 1988; Theeuwes, 1990). In this kind of field of activity, shifting the focus of attention does not refer to initial adoption of a focus from some neutral waiting state (Kahneman et al., 1973). Instead, one re-orients attentional focus to a newly relevant event from a previous state where attention was focused on other data channels or on other cognitive activities (such as diagnostic search, response planning, communication to other agents). We need to understand how some practitioners develop a facility with reorienting attention rapidly to new potentially relevant stimuli (Woods, 1992). Thus, investigating how to aid mode awareness and how to provide cognitive tools to avoid or cope with mode related problems is a fruitful avenue for advancing our more basic understanding of more general issues like the cognitive processes such as control of attention, workload management, mental simulation or, more simply, the panoply of cognitive processes that go under the generic label of situation awareness.

Another design aiding path that has been proposed to deal with mode related problems is forcing functions. Forcing functions are defined as "something that prevents the behavior from continuing until the problem has been corrected" (Lewis and Norman, 1986). Forcing functions can take a variety of forms: the system can prevent the user from expressing impossible intentions ("gag"), it can react to illegal actions by doing nothing, or it can explore with the user what the user's intention was and then help translate this intention into a legal action ("Self-correct", "Teach me", or "Let's Talk About It"). The problem with such forcing functions is that they require (a) that there is only one legal action/strategy for each intention, or (b) that a system is capable of inferring the user's intention to compare it with his input in order to judge the acceptability of the input. Such a system would also have to have access to information on the overall context which can determine whether an action is appropriate. Without these capabilities, it would have to question almost any action – just in case, and run the hazard of over-interrupting at the wrong times.

The last direction is to consider supervisory control of automated resources as a kind of cooperative or distributed multi-agent architecture. One kind of cooperative agent concept would be to support mode awareness as a "management by consent" process which requires that all members of the team, human and machine, need to agree to any input to the system before it is activated. This approach could help a model or trace of all prior system interactions and lead to better prediction of future automated behavior. If automation and

teamwork are supposed to reduce the burden on the practitioner by taking over and sharing tasks, then it seems counterproductive to require that all input is checked and agreed to by every member of the team.

Note that all of the above kinds of recommendation are human-centered in the sense that the costs of clumsy automation are seen in 'human error' and that the avenue for reducing perceived problems in the human element is to recognize that they are symptoms of the complexities produced by the clumsy use of technological possibilities (Woods et al., 1993).

Implications for the Concept and the Study of Situation Awareness

Analyzing the Phenomenon of Situation or Mode Awareness

Our results suggest that the design and the capabilities of advanced automated systems make it more important and at the same time more difficult for their users to maintain awareness of the status and behavior across the different modes of operation of these systems. Despite the fact that vectors of technology change are increasingly challenging mode awareness, little research has yet been done to better understand the relevant human-machine questions. But without this understanding, it will not be possible to develop effective countermeasures to mode related problems. The same deficit can be observed for the issue of situation awareness in general where a long tradition of research has not brought us much closer to being able to understand and support the phenomenon. What kind of research agenda is needed so that the research base can be expanded and practical countermeasures can be developed before technology change creates mode related problems of such magnitude that individual industries cry out for immediate answers?

First, extended efforts to develop the 'right' definition or a consensus definition of situation (and mode) awareness will probably not be constructive. Rather, the term 'situation awareness' should be viewed as just a label for a variety of cognitive processing activities that are critical to dynamic event-driven and multi-task fields of practice. Control of attention (Gopher, 1991), mental simulation (Klein and Crandall, in press), directed attention (Woods, 1992), mental bookkeeping to track the multiple influences that act on a automated dynamic process and the multiple threads of sub-problems and resulting activity to manage them that go on in dynamic incidents are just a few of the cognitive processes that may pass under the label of situation awareness. Analyzing these cognitive processes and understanding what factors affect these processes should be the focus in the attempt to support situation and mode awareness (cf., Endsley, 1988, Tenney et al., 1992, and Sarter and Woods, 1991 for some initial steps). Second, it appears to us to be futile to try to determine the most important contents of situation awareness because the significance and meaning of any data is dependent upon the context in which they appear.

Measuring Situation or Mode Awareness

Conceptual or theoretical developments about the cognitive processes peculiarly associated with supervisory control of dynamic processes are critical if we are to develop

effective measures of mode or situation awareness. Measurement techniques cannot be developed or used in a theoretical vacuum. There are three major categories of measurement, a) subjective ratings, b) explicit performance measures and c) implicit performance measures. The use of subjective measures (e. g., Situation Awareness Rating Technique, SART; Vidulich, 1992), where the operator is expected to rate his or her own level of awareness, is problematic on a variety of grounds, e.g., confusing process and product. It is problematic because there is field data that misassessments color that person's whole standing and recall of the incident evolution (for example, video replay of the participant's behavior coupled with replay of the actual state of affairs is often necessary to get participants to recognize their own misassessments). Subjective measures only seem to make sense when combined with other measurement techniques, for example, in order to learn about how well calibrated were the participants in the evolving incident (extending the concept of how well calibrated is a decision maker to control of attentional focus).

An example of an explicit performance measure to assess situation awareness is Situation Awareness Global Assessment Technique (SAGAT; Endsley, 1988). This method requires that subjects, typically pilots, fly a given mission on a flight simulator. At some random point(s) in time, the simulation is halted and the cockpit displays and outside view is blanked. The pilot is then asked a series of questions about the existing situation. The pilot' s answers are later compared with what was actually going on in the scenario. The agreement between the two serves as a measure of the pilot's situation awareness. The most important problem associated with this technique is that halting the simulation and prompting the pilot for information concerning particular aspects of his situation is likely to disturb the very phenomena which the investigator wishes to observe. All research methods for the study of attentional and awareness related processes suffer from this dilemma -- the methods of observation disturb and change or eliminate the phenomenon under observation. For example, one of the important cognitive constituents of situation awareness is the ability to activate relevant knowledge during the actual process of handling an evolving incident. Prompting the participant for knowledge concerning particular aspects of a situation is itself a kind of retrieval cue and relevance marker that can change what the participant will call to mind. This will reveal what knowledge the pilot can activate when prompted with investigator cues as to relevance, but it will not shed light on what knowledge the participant would activate when alone or see as relevant in a particular situation.

The third approach, implicit performance measures, involves the design of experimental scenarios that include tasks and events that probe the subject's situation awareness (e.g., Sarter and Woods, 1993). In order for this technique to work, the probes have to be operationally significant in the sense that they should provide cues to the operator which if perceived by him should lead to an observable change in behavior. The shortcoming of this technique is that it assumes a direct relationship between situation awareness and performance. This problem can be addressed, in part, however, by means of debriefings in which the attempt is made to determine why a certain behavior did or did not occur. The major advantage of the approach is that it allows for a focused on-line collection of data while trying to minimize the disruptive impact of probes on the behavior of the subject any more than is inevitable in any simulated situation (for a more comprehensive critique of techniques for measuring situation awareness see, Tenney et al., 1992; Sarter and Woods, 1991).

Conclusions

As technology allows for proliferation of more automated modes of operation, human supervisory control faces new challenges. The technological trends create the need for mode awareness – human supervisory controllers tracking what their machine counterparts are doing, what they will do, and why they are doing it. But there is hysteresis in changing training for supervisory controllers and developing displays and interfaces to support collaboration to catch up with the cognitive demands imposed by clumsy use of technological possibilities. The result is evidence from field experiments, incident sampling and accidents that mode related problems in highly automated systems such as loss of mode awareness can contribute to new error forms and new paths towards disasters.

As a consequence, we need to be concerned with the question of how mode awareness can be supported successfully. In order to support mode awareness, as well as situation awareness in general, we need to better understand the set of cognitive processing activities that are involved in these phenomena. In other words, we need to take a process-oriented rather than a product-oriented approach to the analysis of the phenomena of mode and situation awareness. This approach will enable us to identify the reasons for breakdowns in mode and situation awareness, and it will help point the way towards how to train supervisory controllers and how to design cognitive tools that support the monitoring, assessment and awareness demands on supervisory controllers.

References

Adler, P. (1986). New technologies, new skills. *California Management Review*, *29*, 9-28.

Bereiter, S. & Miller, S.M. (1988). Sources of difficulty in troubleshooting automated manufacturing systems. In W. Karwowski, H. Parsaei & M.R. Wilhelm (Eds.), *Ergonomics of Hybrid Automated Systems*. Amsterdam: Elsevier.

Billings, C. E. (1991). *Human-centered aircraft automation philosophy.* (NASA Technical Memorandum 103885). Moffett Field, CA: NASA-Ames Research Center.

Cook, R. I., Potter, S.S., Woods, D.D. & McDonald, J.M. (1991). Evaluating the human engineering of microprocessor-controlled operating room devices. *Journal of Clinical Monitoring* , *7*, 217-226.

Endsley, M.R. (1988, May). *Situation awareness global assessment technique (SAGAT).* Paper presented at the National Aerospace and Electronic Conference (NAECON), Dayton, OH.

Eriksen, C.W. & Murphy, T.D. (1987). Movement of attentional focus across the visual field: A critical look at the evidence. *Perception and Psychophysics, 42 (3)*, 299-305.

Gopher, D. (1991). The skill of attention control: Acquisition and execution of attention strategies. In D. Meyer and S. Kornblum (Eds.), *Attention and Performance XIV*, Hillsdale NJ: Erlbaum.

Jonides, J. & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception and Psychophysics*, *43*, 346-354.

Kahneman, D. (1973).*Attention and Effort.* Englewood Cliffs: Prentice-Hall.

Klein, G. & Crandall, B. (1993). The role of mental simulation in problem solving and decision making. In J. Flach, P. Hancock, J. Caird & K. Vicente (Eds.) *The Ecology of Human-Machine Systems*. Hillsdale NJ: Erlbaum.

Lenorovitz, J.M. (1990). Indian A320 crash probe data show crew improperly configured the aircraft. *Aviation Week & Space Technology, 132(6/25/90)*, 84-85.

Lewis, C. & Norman, D.A. (1986). Designing for error. In D.A. Norman & S.W. Draper (Eds.), *User-centered system design: New perspectives of human-computer interaction*, (pp. 411-432). Hillsdale, NJ: Lawrence Erlbaum Associates.

Moll van Charante, E., Cook, R.I., Woods, D.D., Yue, L., & Howie, M.B. (1992). Human-computer interaction in context: Physician interaction with automated intravenous controllers in the heart room. In IFAC/IFIP/IFOEA/IEA (Eds.), *Proceedings of the Fifth Symposium on Analysis, Design and Evaluation of Man-Machine Systems*. The Hague, Netherlands.

Monk, A. (1986). Mode errors: A user-centered analysis and some preventative measures using key-contingent sound. *International Journal of Man-Machine Studies, 24*, 313-327.

Monnier, A. (1992). *Rapport preliminaire de la commission d'enquete administrative sur l'accident du Mont Sainte Odile du 20 Janvier 1992.*

Norman, D.A. (1981). Categorization of action slips. *Psychological Review, 88(1)*, 1-15.

Norman, D.A. (1988). *The Psychology of Everyday Things.* New York: Basic Books.

Norman, D.A. (1990). The 'problem' of automation: Inappropriate feedback and interaction, not 'over-automation.'*Philosophical Transactions of the Royal Society of London*, B 327.

Sarter, N.B. & Woods, D.D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology, 1(1)*, 45-57.

Sarter, N.B. & Woods, D.D. (1992). Pilot interaction with cockpit automation: Operational Experiences with the Flight Management System. *International Journal of Aviation Psychology, 2(4)*, 303-322.

Sarter, N.B. & Woods, D.D. (1993). *Pilot interaction with cockpit automation II: An experimental study of pilots' model and awareness of the Flight Management System.* Manuscript submitted for publication (Also Cognitive Systems Engineering Laboratory Report CSEL 92-J-11, Columbus, OH: The Ohio State University).

Sellen, A.J., Kurtenbach, G.P., & Buxton, W.A.S. (1992). The prevention of mode errors through sensory feedback. *Human-Computer Interaction, 7*, 141-164.

Tenney, Y.J., Jager Adams, M., Pew, R.W., Huggins, A.W.F., & Rogers, W.H. (1992). *A principled approach to the measurement of situation awareness in commercial aviation.* (NASA Contractor Report No. NAS1-18788). Hampton, VA: NASA Langley Research Center.

Vidulich, M.A. (1992). Measuring situation awareness. In *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 40-41). Santa Monica, CA: Human Factors Society.

Woods, D.D. (1992). *The Alarm Problem and Directed Attention in Dynamic Fault Management.* Cognitive Systems Engineering Laboratory Report CSEL 92-TR-06, Columbus, OH: The Ohio State University.

Woods, D.D. (1993). The price of flexibility. In W. Gray, W. Hefley (Eds.) *Proceedings of the International Workshop on Intelligent User Interfaces*, New York: ACM.

Woods, D.D., Cook, R.I. & Sarter, N.B. (1993) *Clumsy Automation, Practitioner Adaptive Response and 'Human Error.'* Cognitive Systems Engineering Laboratory Report CSEL 92-TR-04, Columbus, OH: The Ohio State University.

The Ubiquitous Three in the Prediction of Situational Awareness: Round Up the Usual Suspects

Thomas R. Carretta , David C. Perry & Malcolm James Ree

United States Air Force, Armstrong Laboratory

Introduction

Aircraft accidents involving a loss of life or aircraft are often blamed on a lack of situational awareness by the pilot or crew. The ability of a pilot to know location in space and in time, as well as keeping track of other aspects of the dynamic environment of aircraft flight, are the common elements of the myriad of definitions of situational awareness (SA). Despite the commonality there is little agreement about what SA is. A survey of recent literature yielded definitions ranging from "a sixth sense" (Hartman & Secrist, 1991; quoting Forrester, 1978) to "mental representations of various flight relevant dimensions" (Andre, Wickens, Moorman, & Boschelli, 1991). However, anecdotal evidence indicates that most pilots consider SA sufficiently important to warrant continued research even without a consensus as to the definition of the trait.

The impetus for this study was provided by the Air Force Chief of Staff. He has directed USAF research laboratories to investigate the human attributes that enable a pilot to develop and maintain SA in combat situations, especially in the F-15, a high performance jet aircraft. Some research issues raised included: a) can SA be measured objectively and reliably; b) can pilot training applicants be screened on SA or on characteristics related to SA; c) can SA be trained; and d) if it can be measured, when and how should this be done?

This prediction study is designed from a personnel selection standpoint. The goal is to develop an understanding of SA and of ways of predicting the amount of SA an applicant possesses or the amount of SA an applicant can learn. The major design features are correlational as opposed to experimental.

The majority of the published literature on SA suffers from a lack of adequate sample size and therefore a lack of statistical power for detecting significant correlational effects. For example, a review of six studies (Andre, Wickens, Moorman, & Boschelli, 1991; Bolstad, 1991; Endsley, 1990; Fracker, 1991; Fracker & Davis, 1991; Kass, Herschler, & Companion,1991) including eight empirical experiments involving SA, showed sample sizes ranging from 8 to 56 with an average sample size of 21.75 subjects. The likelihood of finding significant correlational effects is very low with such small samples. Clearly, larger samples with a higher likelihood of detecting significant effects are required.

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

Both multiple empirical studies and meta-analyses have shown that there are three predictors which are valid for almost all training and job performance criteria. These "ubiquitous" predictors are psychometric *g* (Brand, 1987; Hunter & Hunter, 1984; McHenry Hough, Toquam, Hanson, & Ashworth, 1990; Ree & Earles, 1992), psychomotor skill (Hunter & Hunter, 1984), and the personality construct of "conscientiousness" (Barrick & Mount, 1991). In addition to the Ubiquitous Three, we have included other measures which various theories of human performance (Arthur, Barrett, & Doverspike, 1990; Tett, Jackson, & Rothstein, 1991; Vidulich, 1992) suggest might be useful for prediction of SA. A recent study (Dominguez, Koznik, Gallaway, Whitmore & Armstrong, in press) demonstrated significant differences in information processing speed among different classes of U. S. Air Force pilots. Fighter pilots responded about 10% faster than non-fighter pilots on an aircraft identification task with no loss of accuracy. This result suggests that information processing speed might be predictive of situational awareness, thus measures of information processing speed will be included.

Additionally, from the cognitive realm, measures of speed of working memory, spatial ability, time estimation, and perceptual speed will be included. In the psychomotor realm both general and specific psychomotor skills (Ree & Carretta, 1992) will be assessed. Also, all the personality dimensions which comprise the Big Five (Digman, 1990; Goldberg, 1990; Pedersen, Allan, Laue, Johnson, & Siem, 1991; Tupes & Christal, 1961) will be used. Additionally, operational variables from permanent records such as source of commission (Reserve Officer Training Corps-ROTC, Officer Training School-OTS, Air Force Academy-AFA), Air Force base (AFB) at which initial pilot training took place, and length (months and years) of flying experience will be available for analyses.

The purpose of this study will be to measure traits thought to be related to SA and to use measures of these traits to statistically predict ratings of SA in a concurrent validation design.

Method

Subjects

The subjects will be operational USAF pilots flying the F-15A or F-15C jet fighter aircraft. The F-15A/C were chosen to investigate SA in the air-to-air mission. It is anticipated that there will be between 150 and 200 pilots tested at Eglin, Elmendorf, Kadena, and Langley AFBs. They will be male college graduates ranging in age from about 24 to 45 years old and ranging in military rank from first lieutenant to lieutenant colonel. All will be active duty pilots with from 1 to about 20 years post-pilot-training flying experience.

Measures

Predictors. The predictors of SA for this study will be representatives of the Ubiquitous Three and a variety of other variables which appear to have conceptual or face validity. Tests for which no bibliographic references are provided, are experimental USAF measures.

Cognitive. Multiple studies (Brand, 1987; Earles & Ree, 1991; Jensen, 1992; Kranzler & Jensen, 1991; Kyllonen, in press; Kyllonen & Christal, 1990; Miller & Vernon, 1991; Ree &

Carretta, 1992; Ree & Earles, 1991a, 1992) show that all cognitive tests measure psychometric *g* to some degree regardless of the means of measurement and the outward appearance of the test. For example, Ree and Carretta (1992) found that information processing tests and even psychomotor tests, which are not frequently thought of as being measures of *g*, were about as highly *g* saturated as ordinary paper-and-pencil tests, despite the great dissimilarity of appearance and mode of administration (printed booklets, pencils, and answer sheets versus computer terminals, control sticks, and keypads). See Jensen (1992) for a more complete discussion on the issue of measures of *g*. Although we have listed specific cognitive factors or components for the following tests, it is expected that they will be *g* saturated to some extent.

Air Force Officer Qualifying Test (AFOQT). The AFOQT is a paper-and-pencil multiple aptitude battery. It is composed of several subtests measuring psychometric *g* (Earles & Ree, 1991; Ree & Earles, 1992), and the common factors of verbal, quantitative, spatial, perceptual speed, and aircrew aptitude/interest (Skinner & Ree, 1987). AFOQT subtests are aggregated into composites which are used in commissioning through ROTC and OTS and in the selection of pilot and navigator candidates.

Anticipation. In this velocity estimation test, a target moves from left to right on the screen. When the target reaches line A on the screen, it disappears from view but continues to move at the same velocity. The subject's task is to estimate when the target will cross line B (to the right of line A). Target velocity and point of disappearance vary across test items.

Continuous Opposites. This test measures verbal working memory. Subjects are required to remember the last three words (or their opposites) in a list presented one word at a time. If a word appears in the color red, the subject is required to remember its opposite.

Double Line Smooth Motion. In this velocity estimation test, two lines grow from left to right. The subject's task is to indicate which line will reach the right side of the screen first. The relative velocities of the two lines and the head start given to one of the lines varies across test items.

Four-Term Ordering. This test provides a measure of spatial working memory. Subjects are presented sequentially with three rules regarding the configuration of four block figures. Once all three rules have been reviewed, subjects must choose the complex figure which correctly combines all three rules.

Hartman. This test (Hartman & Secrist, 1991) provides a measure of visual near threshold processing (perceptual speed). Subjects are presented with an image of one of four card suits (hearts, diamonds, spades, or clubs) or a distracter (blank) for a short interval (62, 50, or 33 milliseconds). Subjects are instructed to indicate which of the images was presented.

Instrument Comprehension. In this spatial reasoning test derived from the AFOQT test of the same name (Skinner & Ree, 1987), subjects are presented with a representation of an aircraft compass and horizontal situation indicator. The subject's task is to determine the attitude of the aircraft based on the instrument readings.

Manikin. This test (Benson & Gedye, 1963) provides a measure of spatial transformation. Subjects are presented with an illustration of a man in one of four positions. The image can

be either right side up or upside down and facing toward or facing away. The figure holds a circle in one hand and a square in the other. On each item, the subject is presented with a target object (circle or square) and the figure in one of four positions. The subject's task is to determine which hand (right or left) is holding the target object.

Matrices. In this spatial reasoning test which is similar to the Raven's (1966) Matrices, subjects are presented with an incomplete geometric pattern (the lower right hand corner is missing) and are to choose from several alternatives, which would correctly complete the pattern.

Mental Rotation. This test (Shepard & Metzler, 1971) provides a measure of spatial relations. Subjects are presented sequentially with a pair of letters and are required to make a same or different judgment. The letter pair may be either identical or mirror images, and the letters may either be in the same orientation or rotated in relation to each other. A correct judgment of different occurs when the letters are mirror images of each other regardless of their relative rotation.

Pitch-Roll-Yaw. In this spatial visualization test, representations of two aircraft are displayed side-by-side. The subject must use the right hand control stick and the rudder pedals to maneuver the aircraft on the right to match the aircraft on the left (target) in the pitch, roll, and yaw axes.

Rapid Serial Classification: 4-Square. This test measures spatial reasoning ability. Subjects are shown a 4-square (2-by-2) display in which a letter pattern can be drawn (C, X, or Z) between points. Subjects must determine which letter is being drawn by following the pattern of dots as they are sequentially illuminated and extinguished.

Scheduling 2. In this measure of divided attention, five horizontal logarithmic scales can be presented. A line beneath each scale increases at a unique, constant rate. Each line and scale appears on a separate screen which may be viewed by entering the scale number on the response keypad. Subjects score points equal to the current value of the line displayed on the scale by pressing the ENABLE key. When the ENABLE key is pressed, the subject's total score is incremented by the value of the line which is then reset to 0, where it will start increasing again. If the value of a line reaches the upper limit of the scale, and the subject has not responded by pressing the ENABLE key, the value of the line will reset to 0 without the subject receiving any points.

Simultaneous Figure Matching. In this spatial reasoning test (Palmer, 1977) subjects view two geometric line drawings formed by connecting dots on two 3-by-3 dot matrices. The subject's task is to determine whether the two line drawings are identical.

Single Line Smooth Motion. In this velocity estimation test, a single line grows from left to right on the screen. The line disappears from view before reaching the right side of the screen. Subjects must estimate when the line would have reached the right side of the screen, had it continued at the same rate.

Spatial Orientation. In this test of spatial orientation, subjects are presented with a cube that has a geometric figure inside of it. Subjects are instructed to imagine that they are

viewing the geometric figure through one of several "windows" in the cube. The subject's task is to determine what the geometric figure would look like when viewed through a particular "window." In the second part of the task, subjects are shown the geometric figure and must determine through which "window" it is being viewed.

Three-Dimensional Mental Rotation. In this spatial relations (Shepard & Metzler, 1971) test, a pair of geometric figures is presented side-by-side on the screen. Each figure is composed of 10 equally sized cubes, connected together so as to form two right angles with itself. Each figure pair may be either identical or mirror images, and the figures may be either in the same orientation or rotated in relation to each other. Subjects must determine whether the figure pairs are identical (except for orientation) or mirror images.

Verification Span. This test modeled after the Daneman and Carpenter (1980) reading span test, measures verbal working memory. Subjects must memorize an individually presented word, while simultaneously responding "true/false" to declaratory statements (e.g., San Antonio is the capital of Texas). Subjects are allowed eight seconds to respond to each statement. Then the next word to be memorized/statement combination is presented. After the list of words to be memorized and statements have been presented, subjects must recall the memorized words in the proper order. Progressively longer lists of word/statement combinations are used.

XYZ Assignment: Synthesis Add/Subtract. This test provides a measure of spatial working memory. Subjects are required to combine or delete simple line figures assigned to three letters (X, Y, and Z). Two figures are assigned to each letter in the form of an addition or subtraction equation. Subjects must mentally combine or delete the lines of these figures and then memorize the combination. Information about one figure is sometimes needed to solve the equation for one of the other figures.

Psychomotor

Complex Coordination. This test provides a measure of multilimb coordination (Fleishman, 1964). Using a dual-axis right-hand control stick, subjects are required to keep a 1" cross centered on a dotted line cross that bisects the screen horizontally and vertically. Simultaneously, using the left-hand single-axis control stick, subjects have to keep a 1" vertical bar horizontally centered at the base of the screen.

Laser Aiming Task 1. This test provides a measure of multilimb coordination and aiming. Subjects maneuver left and right foot pedals to aim a "laser gun" at aircraft which move horizontally across the screen. Subjects fire the" laser gun" by pressing the ENABLE key. Speed, distance, and direction (left or right) of the target aircraft vary across trials.

Laser Aiming Task 2. This test is similar to Laser Aiming Task 1, except that subjects are instructed to imagine they are shooting from an aircraft at the bottom of the screen. Subjects must match the apparent altitude (size) of the target and the "laser gun" to get the laser beam on target.

Scanning and Allocating. This compensatory tracking task provides a measure of control precision (Fleishman, 1964). In this test, subjects are required to simultaneously maintain the

vertical alignment of four vertical lines using the right-hand control stick. Subjects can control only one vertical line at a time, switching among the lines by using the numeric keypad.

Time Sharing 2. This test provides measures of attention and the psychomotor factors of reaction time and rate control (Fleishman, 1964). The first part of the test involves learning a compensatory tracking task, where subjects maneuver the right-hand control stick to keep a "gunsight" centered on an airplane. The second part of the test involves learning an attention task. Numbers appear one at a time in sequence at the lower part of the screen. The number sequence is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 0, 1, 2, 3, etc. Occasionally, a number will be missing from the sequence (e.g., 0, 1, 2, 3, 5, 6, 7, ... [4 is missing]). Subjects are required to type the missing number on the keypad. During the final part of this test, subjects simultaneously perform tracking and attention tasks.

Personality

Test items are combined to create measures of the Big Five (Goldberg, 1990) personality dimensions of: Extraversion, Emotional Stability, Agreeableness, Conscientiousness, and Openness to Experience. Subjects are presented a list of 205 adjectives and must indicate the extent to which each adjective is self-descriptive from "not at all characteristic of me" to "extremely characteristic of me."

Experience. It is necessary to take experience into account to understand the relationships between the predictors and the criteria. A complete flight history will be obtained including the number of hours for each aircraft flown and the type of qualification achieved (e.g. wingman, 2-ship flight lead, 4-ship flight lead, instructor pilot).

Criteria. The criteria will be supervisory, self, and peer SA ratings collected at the operational sites. These rating scales were developed (Houck, Whitaker, & Kendall, 1991) from task analyses conducted by psychologists using experienced F-15 pilots as subject matter experts (SMEs). These pilot SMEs identified tasks essential to air combat success and required for SA (W. Waag, personal communication, January 6, 1993). This resulted in 31 items considered to be the salient personal traits and job tasks related to SA within an air combat environment. Rating items represent general traits (e.g. reasoning, spatial ability), tactical game plan (e.g. tactics development, execution), systems operation (e.g. weapons systems proficiency), communication (e.g. effectiveness), information interpretation (e.g. threat prioritization), and tactical employment (e.g. targeting weapons, threat evaluation, mutual support). Standardized definitions for each of the items will be provided to every rater to establish consistency in meanings. Each of the 31 items is rated on a six point Likert scale from 1- "Acceptable" to 6- "Outstanding." Supervisors such as squadron commanders, operations officers, assistant operations officers, and flight commanders will rate pilots on the 31-item SA criterion instrument. Other raters include the squadron weapons officer and the flight examiner. Almost all subjects will be rated by multiple raters. Supervisors' ratings reflect their perceptions of each pilot compared to other F-15 pilots.

Supervisors will provide an index of familiarity for each pilot rated. The index consists of a self report of the number of sorties flown with the pilot and an estimate, on a Likert scale

(1-"Inadequate Knowledge" to 6-"Comprehensive Knowledge"), of the raters ability to evaluate the pilot's overall fighter aircraft skills.

Each pilot will make self ratings in the same fashion as the supervisory ratings and using the same rating 31 items and definitions. An example of the supervisory and self rating scale is provided in the Appendix.

For peer ratings, pilots will rate other pilots in their squadron with whom they have flown. Overall fighter ability and SA ability will be scored on a six-point Likert scale ranging from 1- "Acceptable" to 6- "Outstanding." Pilots will then rank order peers rated from 1- "the best I've flown with" to N, the number of peers rated, indicating their standing on the trait of SA. The global rating questions are found in the appendix. Pilots will not provide peer ratings for other pilots about whom they have insufficient knowledge.

Weights by which to multiply each criterion item have been established. These were the mean rating of importance for SA for each of the 31 items estimated by all squadron commanders, operations officers, and assistant operations officers. In addition, other approaches to weighting have been proposed and will be evaluated. Weights are intended for use in supervisory and self ratings.

Criterion Issues. Several issues pertaining to the criteria remain unresolved. The first of these is whether to use multiple criteria or one criterion (see Campbell, McHenry, & Wise 1990). If one criterion is desired, there are several ways that the individual criteria (supervisory, self, and peer ratings) could be combined. For example, the first principal component of the criteria will, *per force,* yield the most reliable and likely the most predictable criterion composite. If multiple criteria are acceptable, then each may be regressed on the predictors and evaluated separately leaving the determination of importance to the judgment of managers. Further, canonical correlations of the predictor set and criteria could be used to evaluate the relationship between the two sets of variables.

Within each of the criteria there can be both global and specific ratings which are amenable to the same treatment described for sources of the ratings (supervisory, self, and peer). For example, the final global rating could be regressed on the predictors as a single criterion or the first principal component (and possibly others, depending on the amount of variance for which they account) could be regressed on the predictors. Alternately a factor analysis could disclose several meaningful factors, each to be used as criteria.

A second issue is estimation of the reliability of the criteria. The maximum correlation between any two variables is limited by the product of the square roots of their reliabilities. (Hunter, Schmidt, & Jackson, 1982). To understand the true correlation of the predictors with the criteria, it is necessary to adjust the observed correlations for reliability (Carretta & Ree, in press). The research design permits the direct estimation of reliability through inter-rater reliability via intra-class correlation when there are multiple raters. Reliability will be indirectly estimated using communality estimates based on squared multiple correlations as frequently done in factor analyses. The communality is the *lower bound* estimate of the reliability as it does not include the reliable variance attributable to specific variance (Baggaley, 1964).

A third issue is how to scale the criterion ratings to correct for difference in metrics of the raters. The most promising procedure is the use of categorical variables representing the raters which will correct for the mean difference in ratings. Corrections for the variability of the raters are currently under study.

A final issue is the use of the importance weights. While these weights are intuitively appealing, Wilks' (1938) theorem argues against their use. Wilks provided a proof that the

correlation of two linear composites (composite = _ weights x item ratings) approaches 1.0 under the common conditions of all positive correlation among the variables (items) and a sufficient number of variables. This is likely to be the case in the criterion ratings. The utility of the importance weights will be studied.

Apparatus

All of the predictor measures except the AFOQT will be administered by the computer based apparatus. The apparatus consists of a microcomputer and monitor built into a carrel designed to minimize distractions. Subjects will respond to the tests by manipulating individually or in combination, a dual-axis control stick on the right side, a single-axis control stick on the left side, a mouse, foot pedals, and a customized keypad. The keypad includes keys numbered 0 to 9, an ENABLE key in the center, and a bottom row with Yes and No keys, and others for same/different responses (S/D) and left/right responses (L/R).

Procedure

Subjects will be tested on the computerized battery at their duty locations (operational flying wings). Supervisory, self, and peer ratings of SA will be collected independently.

The test battery used in this study requires about five hours to complete. In order to minimize mental and physical fatigue, the test battery will be administered in two 2.5 hour sessions scheduled on different days. Order of the test batteries and the tests within each battery will be randomized.

The AFOQT was administered for operational qualification for the OTS and ROTC commissioning sources. In some cases, the AFOQT may have been taken several years prior to the other predictors.

Analyses

Because there are many predictors and relatively few subjects, some methodology for reducing the number of predictors is necessary. This is expected to be accomplished by the use of factor or principal components analyses dependent on the magnitude and direction of the intercorrelations of the predictors (Gulliksen, 1950; Wilks, 1938).

The issue of single versus multiple criteria will be addressed by investigation of a) the factor structure of the criteria and b) the rank correlation of individuals on the single criterion and the multiple criteria. If the factor structure of the multiple criteria shows a large first factor and very small succeeding factors, a single (or composite) criterion may embody all the reliable variance. If the rank correlations of the criteria approach 1.0, corrected for attenuation due to unreliability, each criterion would rank the subjects about the same and multiple criteria would not be necessary. This issue must be addressed separately for each rating source.

Criterion reliability will be derived through communality estimation (Baggaley, 1964). This requires that each criterion measure be regressed on all the other criterion measures. The resulting R^2 is the estimate of reliability. Reliability of the predictors will be estimated by

internal consistency when assumptions are met and by communality when the assumptions of measures of internal consistency are not met.

A linear models approach using both continuous and categorical variables will be used to correct the metrics of the criteria provided by various raters. The categorical variables will identify supervisor making the rating and the base where the rating was made. The continuous variables will be the test scores or aggregates of test scores and experience. This is analogous to the procedure used by Ree and Earles (1991b) to correct for mean differences in technical training grades across many schools. Various models will be investigated by imposing restrictions which set equal the rater adjustment coefficients to the categorical variables. The smallest number of categorical variables which adequately represent the raters and correct the means of the ratings will be used.

Tests of linear regression models will be the validity paradigm of the analyses. In each statistical test, there will be two regression equations. One, the full model, contains all the variables of interest and the other, the reduced model, contains a subset of the variables omitting the variable or variables whose incremental predictiveness is being evaluated (Ward & Jennings, 1973). A step-down hierarchical approach will be used where the full model (with the largest number of predictors) will be tested against the next fullest model and so on. This is not a stepwise regression procedure which would be ill-advised in a range restricted sample. The linear regression models will be specified to answer questions of interest. All statistical tests will be conducted at the $p < .05$ Type I error rate. If need be, the experiment-wise error rate (i.e. the accumulation of Type I error across statistical hypotheses) will be controlled by the use of the Bonferroni inequality (Miller, 1966).

Summary

Previous studies of SA have suffered from a lack of a theoretical foundation, small sample sizes, and a deficiency of empirical validity. We have used the theoretical basis of the Ubiquitous Three as a foundation because they have been shown to be predictive of numerous criteria. Predictor and criterion data will be collected at several locations and analyzed to establish predictive relationships. Results from this study will guide our future research in the prediction of situational awareness.

References

Andre, A. D., Wickens, C. D., Moorman, L., & Boschelli, M. M. (1991). Display formatting techniques for improving situation awareness in the aircraft cockpit. *The International Journal of Aviation Psychology, 1,* 205-218.

Arthur, W., Barrett, G. V., & Doverspike, D. (1990). The validation of an information-processing-based test battery for the prediction of handling accidents among petroleum-product transport drivers. *Journal of Applied Psychology, 75,* 621-628.

Baggaley, A. R. (1964). *Intermediate Correlational Methods*. New York: John Wiley & Sons.
Barrick, M. R., & Mount, M. K. (1991. The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1-26.

Benson, A. J., & Gedye, J. L. (1963). *Logical process in the resolution of orientation conflict* (Report 259). Farnborough, UK: Royal Air Force, Institute of Aviation Medicine.

Bolstad, C. A. (1991). Individual pilot differences related to situation awareness. *Proceedings of the Human Factors Society, 35th annual meeting* (pp. 52-56). San Francisco, CA.

Brand, C. (1987). The importance of general intelligence. In S. Modgil & C. Modgil (Eds.) *Arthur Jensen: Consensus and controversy*. New York: The Falmer Press.

Campbell, J. P., McHenry, J. J., & Wise, L. L. (1990). Modeling job performance in a population of jobs. *Personnel Psychology, 43,* 313-333.

Carretta, T. R., & Ree, M. J. (in press). Basic Attributes Test (BAT): Psychometric equating of a computer-based test.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal *of Verbal Learning and Verbal Behavior, 19,* 450-466.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. In M. R. Rosenzwieg & L. W. Porter (Eds.). *Annual Review of Psychology, 41,* 417-440

Dominguez, C. O., Kosnik, W. D., Gallaway, R. A., Whitmore, J.,& Armstrong, S. (in press). *Assessing automatic and near-threshold processing abilities as components of situational awareness in U. S. Air Force pilots.* A paper submitted to the 37th Annual meeting of the Human Factors and Ergonomic Society, Seattle, WA.

Earles, J. A., & Ree, M. J. (1991). *Air Force Officer Qualifying Test: Estimating the general ability component* (AL-TR-1991-0039). Brooks AFB, TX: Armstrong Laboratory, Human Resources Directorate, Manpower and Personnel Division .

Endsley, M. R. (1990). Predictive utility of an objective measure of situation awareness. *Proceedings of the Human Factors Society, 34th annual meeting,* (pp. 41-45), Orlando, FL.

Fleishman, E. A. (1964). *The structure and measurement of physical fitness.* Englewood Cliffs, NJ: Prentice-Hall.

Forrester, L. (1978). *Fly for your life.* (2nd ed.) New York: Bantam.

Fracker, M. L. (1991). *Measures of situation awareness: An experimental evaluation* (AL-TR-1991-0127). Wright-Patterson AFB, OH: Human Engineering Division, Crew Systems Directorate.

Fracker, M. L., & Davis, S. A. (1991). Explicit, implicit, and subjective rating measures of situation awareness in a monitoring task (AL-TR-1991-0091). Wright-Patterson AFB, OH: Human Engineering Division, Crew Systems Directorate.

Goldberg, L. R. (1990). An alternate "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology, 59,* 1216-1229.

Gulliksen, H. (1950). *Theory of Mental Tests.* New York: John Wiley & Sons.

Hartman, B. O., & Secrist, G. E. (1991). Situational awareness is more than exceptional vision. *Aviation, Space, & Environmental Medicine, 62,* 1084-1089.

Houck, M. R., Whitaker, L. A., & Kendall, R. R. (1991). *Behavioral taxonomy for air combat: F-15 defensive counter-air mission* (UDR-TR-91-147). Dayton OH: University of Dayton Research Institute.

Hunter, J. E., & Hunter, R. (1984). Validity and utility of alternate predictors of job performance. *Psychological Bulletin, 96,* 72-98.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-Analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.

Jensen, A. R. (1992). Commentary: Vehicles of g. *Psychological Science, 3,* 275-278.

Kass, S. J., Herschler, D. A., & Companion, M. A. (1991). Training situational awareness through pattern recognition in a battlefield environment. *Military Psychology ,3,* 105-112.

Kranzler, J. H., & Jensen, A. R. (1991). The nature of psychometric *g* : Unitary process or a number of independent processes? *Intelligence, 15,* 397-422.

Kyllonen, P. C. (in press). Aptitude testing inspired by information processing: A test of the four-sources model.

Kyllonen, P. C. , & Christal, R. E. (1990) Reasoning ability is (little more than) working memory capacity?! *Intelligence, 14,* 389-433.

McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology, 43,* 335-354.

Miller, R. G. (1966). *Simultaneous statistical inference.* New York: McGraw-Hill.

Miller, L. T., & Vernon, P. A. (1991). The general factor in short-term memory, intelligence, and reaction time. *Intelligence, 16,* 5-29.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology, 9,* 441-474.

Pedersen, L. A., Allan, K. E., Laue, F. J., Johnson, J. R., & Siem, F. M. (1992). *Personality theory for aircrew selection and classification* (AL-TR-1992-0021). Brooks AFB, TX: Armstrong Laboratory, Human Resources Directorate, Manpower and Personnel Research Division.

Raven, J. C. (1966). *Advanced Progressive Matrices.* New York: Psychological Corporation.

Ree, M. J., & Carretta, T. R. (1992). *The correlation of cognitive and psychomotor tests* (Al-TP-1992-0037). Brooks AFB, TX: Armstrong Laboratory, Human Resources Directorate, Manpower and Personnel Research Division.

Ree, M. J., & Earles, J. A. (1991a). The stability of *g* across different methods of estimation. *Intelligence, 15,* 271-278.

Ree, M. J., & Earles, J. A. (1991b). Predicting training success: Not much more than *g*. *Personnel Psychology, 44,* 321-332.

Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science, 1,* 86-89.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701-703.

Skinner, J., & Ree, M. J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and factor analyses of Form O.* (AFHRL-TR-86-68). Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44,* 703-742.

Tupes, E. C., & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings* (ASD-TR-61-97). Lackland AFB, TX: Personnel Laboratory, Aeronautical Systems Division

Vidulich, M. (1992). *Cognitive and performance components of situation awareness: SAINT Team task one report.* Unpublished manuscript.

Ward, J. H., Jr., & Jennings, E. (1973). *Introduction to linear models.* Englewood Cliffs, NJ: Prentice-Hall.

Wilks, S. S. (1938). Weighting systems for linear functions of correlated variables when there is no dependent variable. *Psychometrika,3.*23-40.

Appendix

Situational awareness supervisory and self rating form

SITUATION AWARENESS RATING SCALE

Rater ID #:____________ Pilot ID #:____________

Relative Ability Compared With Other F-15C Pilots						
Item Ratings	**Acceptable**		**Good**		**Outstanding**	
General Traits	**1**	**2**	**3**	**4**	**5**	**6**
1. Discipline	_	_	_	_	_	_
2. Decisiveness	_	_	_	_	_	_
3. Tactical knowledge	_	_	_	_	_	_
4. Time-sharing ability	_	_	_	_	_	_
5. Reasoning ability	_	_	_	_	_	_
6. Spatial ability	_	_	_	_	_	_
7. Flight management	_	_	_	_	_	_
Tactical Game Plan						
8. Developing plan	_	_	_	_	_	_
9. Executing plan	_	_	_	_	_	_
10. Adjusting plan on-the-fly	_	_	_	_	_	_
System Operation						
11. Radar	_	_	_	_	_	_
12. TEWS	_	_	_	_	_	_
13. Overall weapons system proficiency	_	_	_	_	_	_
Communication						
14. Quality (brevity, accuracy, timeliness, completeness)	_	_	_	_	_	_
15. Ability to effectively use comm information	_	_	_	_	_	_
Information Interpretation						
16. Interpreting VSD	_	_	_	_	_	_
17. Interpreting RWR	_	_	_	_	_	_
18. Ability to effectively use AWACS/GCI	_	_	_	_	_	_
19. Integrating overall information (cockpit displays, wingman comm, controller comm)	_	_	_	_	_	_
20. Radar sorting	_	_	_	_	_	_
21. Analyzing engagement geometry	_	_	_	_	_	_
22. Threat prioritization	_	_	_	_	_	_
Tactical Employment-BVR Weapons						
23. Targeting decisions	_	_	_	_	_	_
24. Fire-point selection	_	_	_	_	_	_
Tactical Employment - Visual Maneuvering						
25. Maintain track of bogeys/friendlies	_	_	_	_	_	_
26. Threat evaluation	_	_	_	_	_	_
27. Weapons employment	_	_	_	_	_	_

Tactical Employment - General

28. Assessing offensiveness/defensiveness	_	_	_	_	_	_
29. Lookout (VSD interpretation, RWR monitoring, visual lookout)	_	_	_	_	_	_
30. Defensive reaction (chaff, flares, maneuvering, etc.)	_	_	_	_	_	_
31. Mutual support	_	_	_	_	_	_

Overall Situational Awareness[3]	_	_	_	_	_	_
Overall Fighter Ability	_	_	_	_	_	_

---

[3] Items 1 through 31 are used for both supervisory and self ratings. The overall fighter ability and SA items are completed by peers.

Comparison of Pilots' Acceptance and Spatial Awareness When Using EFIS vs. Pictorial Display Formats for Complex, Curved Landing Approaches

Anthony M. Busquets[1], Russell V. Parrish[1], Steven P. Williams[2] & Dean E. Nold[3]

[1] Cockpit Technology Branch, Flight Management Division
[2] Joint Research Projects Office, AMSEL-RD-C2-TJ, CECOM, U.S. Army
[3] Purdue University

Introduction

Advances in next-generation cockpits are being made possible by the rapid progress in display media, graphics and pictorial displays, and computer technologies, along with human factors methodologies. These technologies conceivably may enable the design of cockpits contributing to improved crew situation awareness, safety, operational efficiency and reduced crew workload during critical mission phases (Hatfield & Parish, 1990). In pursuing these advancements, research programs have been established by government and industry to develop and exploit these technologies. One such program involves the use of "synthetic vision" to enable transport operations under restricted-visibility conditions as well as provide the cornerstone technology for more advanced aircraft, such as a high-speed civil transport that, because of the complex aerodynamic and economic requirements, may have limited forward visibility.

Various studies have been undertaken to assess the requirements (Regal & Whittington, 1993) and to determine the performance (*Proceedings of the 7th Plenary Session of the Synthetic Vision Certification Issues Study Team,* 1992) of synthetic vision systems. One study (Swink & Goins, 1992) has indicated numerous potential benefits for a future high-speed civil transport (HSCT) in which synthetic vision is used in lieu of drooping the nose for landing, taxi, take-off. These potential benefits include improved aerodynamic efficiency, reduced weight, and as much as 15% reduction in take-off gross weight (TOGF) through reduced fuel reserves. Synthetic vision capabilities are defined herein as the resourceful merging of imaging sensors (such as fog cutting sensors), pictorial graphics displays, and advanced navigational aids (such as Differential Global Positioning System). It is also generally accepted that there is an ever increasing interrelationship between onboard capabilities and airspace management systems, and that, therefore, higher levels of crew situation awareness are required in order to improve performance and safety (Brahney, 1992). Initial investigations are being conducted on cockpit flight displays aimed at optimizing the *spatial* awareness component of situation awareness (Dorighi & Ellis, 1991; Dorighi, Grunwald, &

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

Ellis, 1992; Regal, Rogers, & Boucek, 1988). This paper will focus on large-screen, integrated, pictorial displays as an approach to synthetic vision technology and the problem of optimizing crew spatial awareness.

To understand situation awareness (SA) in the context of commercial transport operations, which is the focus of this research, a definition is necessary. Regal, et al. (1988), states that SA implies "that the pilot has an integrated understanding of the factors that will contribute to the *safe* flying of the aircraft under normal or non-normal conditions". As SA increases, "the pilot is increasingly able to 'think ahead of the aircraft', and that he can do this for a wider variety of situations." This entails "a knowledge of present states, future goals, and the procedures used to get from one to the other." Regal goes on to expound that, for the commercial pilot, another dimension of SA lies in its being defined by a number of individual components.

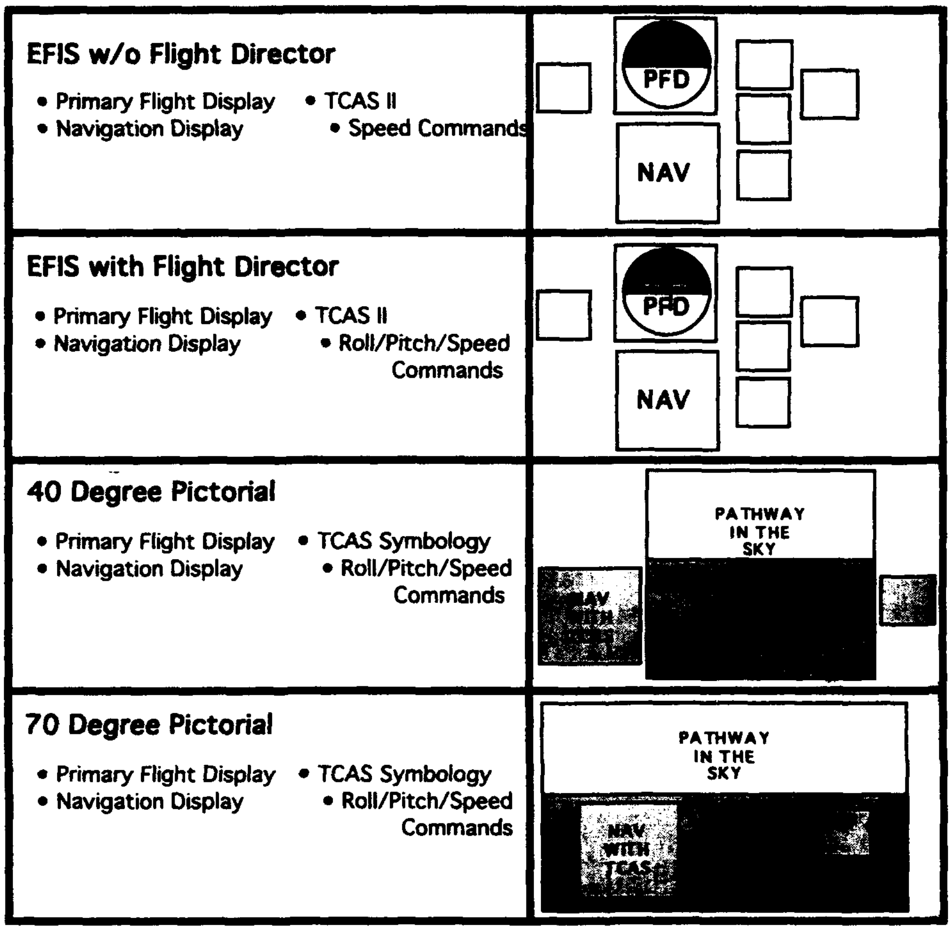


Figure 1. Spatial awareness simulation study display formats. Objective: Compare pilots' spatial awareness using "All Glass" transport pilots with various displays.

The objective of this experiment was to evaluate and compare the *spatial* awareness component of pilots using displays representative of conventional Electronic Flight Information Systems (EFIS) to two wide-field-of-view pictorial display concepts (Figure 1). Two formats, exemplary of a Boeing-757 layout of instrumentation, were used as the representative conventional EFIS formats. In the four alternate display concepts which were compared, the EFIS formats, used as baselines, were identical with the exception that one incorporated a flight director (with commands displayed on two perpendicular needles in the attitude display), while the other forced the pilot to fly raw deviation error (ILS localizer and glideslope indicators), without the benefit of flight director guidance. Both formats were included for calibration purposes as it would seem that spatial awareness would be quite different for the two conditions. That is, concentrating on centering the flight director needles might be expected to reduce the pilot's awareness of surrounding events, while flying raw position errors might increase his spatial awareness. The two pictorial concepts were identical "pathway-in-the-sky" formats, varying only in horizontal field-of-view (40 and 70 degree presentations). Further explanation of the display formats will follow in the section on Display Conditions.

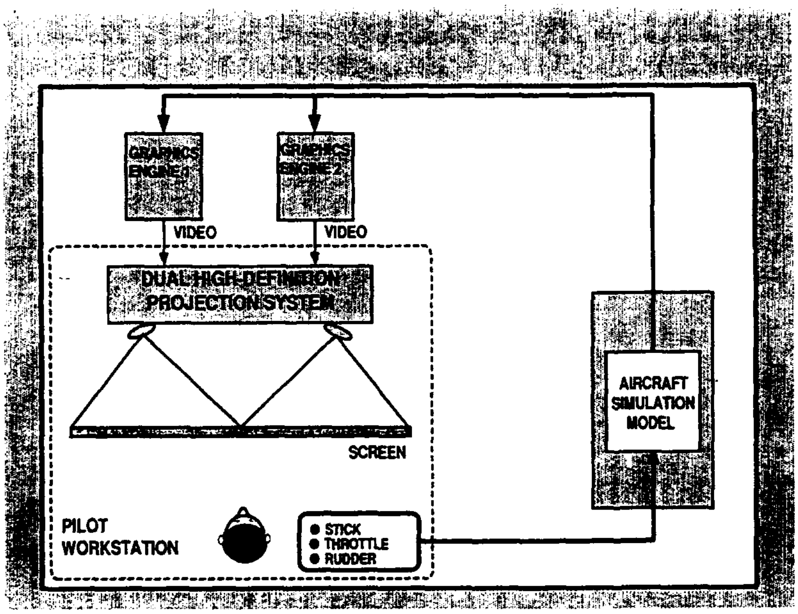


Figure 2. VISTAS architecture.

Simulator Description

The Cockpit Technology Branch at Langley Research Center has developed a flexible, large-screen flight display research system, named VISTAS (Visual Imaging Simulator for Transport Aircraft Systems), which was utilized to carry out this experiment. The simulator is comprised of the following elements: visual system hardware, graphics generation hardware and software, an aircraft mathematical model, and computer implementation (Figure 2). The visual and interactive control elements of this flight display research tool have been integrated as a piloted workstation in order to explore the advantages and limitations of large-screen, pictorial, reconfigurable display concepts and associated interactive techniques.

Simulator Visual System

The core of the visual system is embodied in dual, full-color, high-resolution CRT projectors that are configured to vary the projected display's aspect ratio by edge-matching and overlapping the images from each projector. Since each projected image is 15 inches in height by 20 inches in width (standard 3:4 aspect ratio), a maximum 15 by 40 inch image can be achieved. The images are generated by the dual graphics display generators operating in conjunction, utilizing the same visual database in order to produce a single, large-screen, integrated picture (combined by the projection system onto the rear-projection screen that serves as the simulated aircraft's main instrument panel). Each generator provides image resolutions up to 1280 x 1024 pixels in a 60 Hz progressive scan format (per projector). Given that the design-eye reference point (DERP) for transport cockpit applications is typically around 28 inches, the full 40 inch wide display provides a maximum 70 degree field-of-view (FOV).

Aircraft Mathematical Model and Computer Implementation

A simplified six-degree-of-freedom mathematical model of a transport aircraft was used in this study to provide the interaction between the pilot and the flight display formats. The linear transfer functions and gains were obtained empirically to represent a fixed-wing generic transport aircraft. Turbulence was introduced into the mathematical model through the addition of gust components to the body-axis longitudinal and lateral velocity variables. The level of turbulence was considered to be moderate by the participating pilots.

Simulator Cockpit

The pilot workstation was configured as the pilot side of a generic transport, fixed-wing aircraft with the pilot's seat designed to position the subjects so that their eyes were at the DERP. The workstation also accommodated the dual-head projection system and the rear-projection screen that simulated the instrument panel. Pitch and roll inputs to the aircraft mathematical model were provided in the workstation by a two-degree-of-freedom sidearm

hand-controller with spring-centering. Throttle inputs were provided by a throttle lever that utilized a voltage-referenced potentiometer as the signal source. Typical self-centering rudder pedals provided yaw inputs.

The display screen (instrument panel) was tilted so as to provide a 17 degree line-of-sight (from horizontal) over the top of the screen, which is typical of over-the-glareshield views in most aircraft. The screen's display surface was set perpendicular to the pilot's line-of-sight.

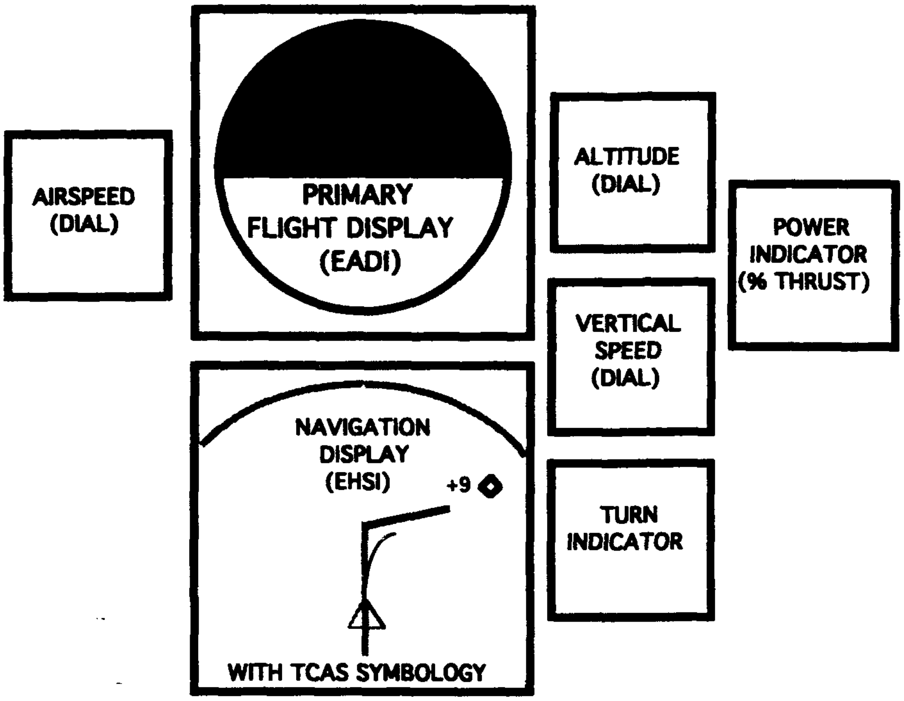


Figure 3. Over-and-under arrangement of a conventional Primary Flight Display and Navigation Display with supporting instrumentation.

Display Conditions

This experiment attempted to assess the "spatial" awareness component of SA of pilots while using integrated pictorial displays as compared to conventional EFIS formats. The two EFIS displays, utilized as baseline measures, differed only in that one lacked the flight director "bars." The basic instrument arrangement was a T-arrangement with a B-757 duplicate of a Primary Flight Display (PFD) over a Navigation Display (ND) (Figure 3). To the left of the PFD was a typical airspeed indicator dial and to the right were typical altitude, vertical speed, and turn coordinator instruments arranged over one another in that order. Non-standard (but used in all four display concepts) was a power indicator which integrated

engine and ambient information to display actual power (including engine spool-up) in percent thrust (Abbott, 1989). Also presented on the power indicator was power commanded by the throttle setting and power desired by the Flight Management System (FMS) for flying the programmed approach.

For the integrated pictorial display formats, a computer-generated out-the-window (OTW) view, with overlaid Head-up Display (HUD) symbology, was presented (Figure 4).

One pictorial concept was rendered in a 70 degree FOV format and the other in a 40 degree FOV format (Figure 1). The OTW portion of the display consisted of a pathway-based approach, depicted by green goalposts whose width and height corresponded to fractions of lateral and vertical ILS beam errors (1/4 and 1/2 dots respectively). Also, a tiled roadway consisting of 20 tiles was presented within the goalposts to aid in vertical station-keeping and to present a speed cue. The HUD symbology included airspeed and altitude vertical tapes, roll and pitch scales (in degrees), as well as a horizontal heading tape. All of the tapes incorporated Flight Management System (FMS) command "bugs." The heading tape also showed ground track while the airspeed tape also showed groundspeed. A vertical speed indicator was integrated onto the altitude tape as a growing/shrinking barber pole with a digital vertical speed tag (whose position on the altitude scale would denote the altitude to be attained in one minute based on current vertical speed). The central HUD symbology consisted of a diamond, depicting pitch attitude, and "waterline" symbols for instantaneous and predicted flight path vectors. The display was attitude-centered with rate command control, although the pilots attempted to control the flight path vector. A secondary "smoked-glass" (see-through) Navigation Display (ND) was presented on the left side of the pictorial displays, basically duplicating the EFIS Navigation Display. Thus, horizontal situation display information was provided that also depicted traffic within the OTW display FOV (delineated by the acute lines about the ownship centerline) as well as traffic outside the FOV.

Table 1.

□	UNFILLED BLUE SQUARE	TCAS OFF (UNDER 500 FEET)
◇	UNFILLED BLUE DIAMOND	<u>NON-THREAT</u>
◆	SOLID BLUE DIAMOND	<u>PROXIMITY TRAFFIC</u> NON-THREATENING WITHIN 1,200 FT ALTITUDE AND 6 NMI RANGE
○	SOLID YELLOW CIRCLE	<u>TRAFFIC ADVISORY</u> WITHIN 1,200 FT ALTITUDE AND TIME =< 45 SECONDS
■	SOLID RED BOX	<u>RESOLUTION ADVISORY</u> ESTIMATED MISS DISTANCE =< 750 FT AND TIME =< 30 SECONDS

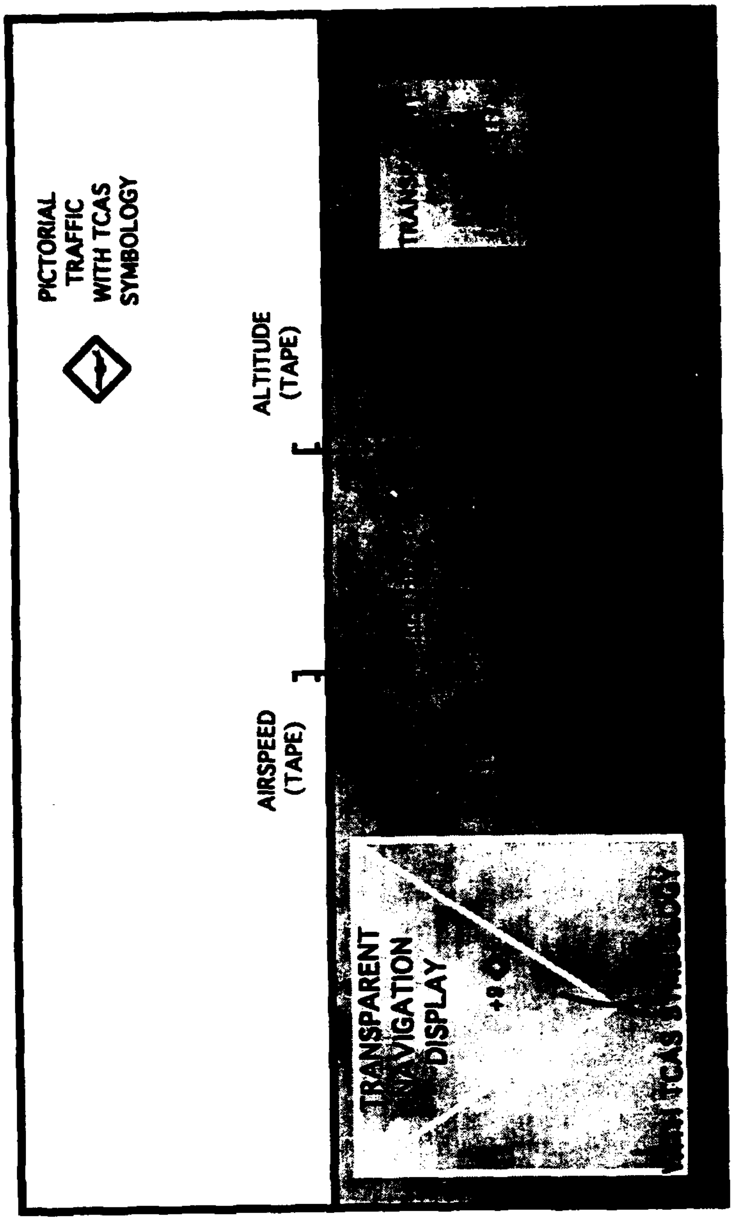


Figure 4. Seventy degree field-of-view, large screen, integrated, pictorial display concept.

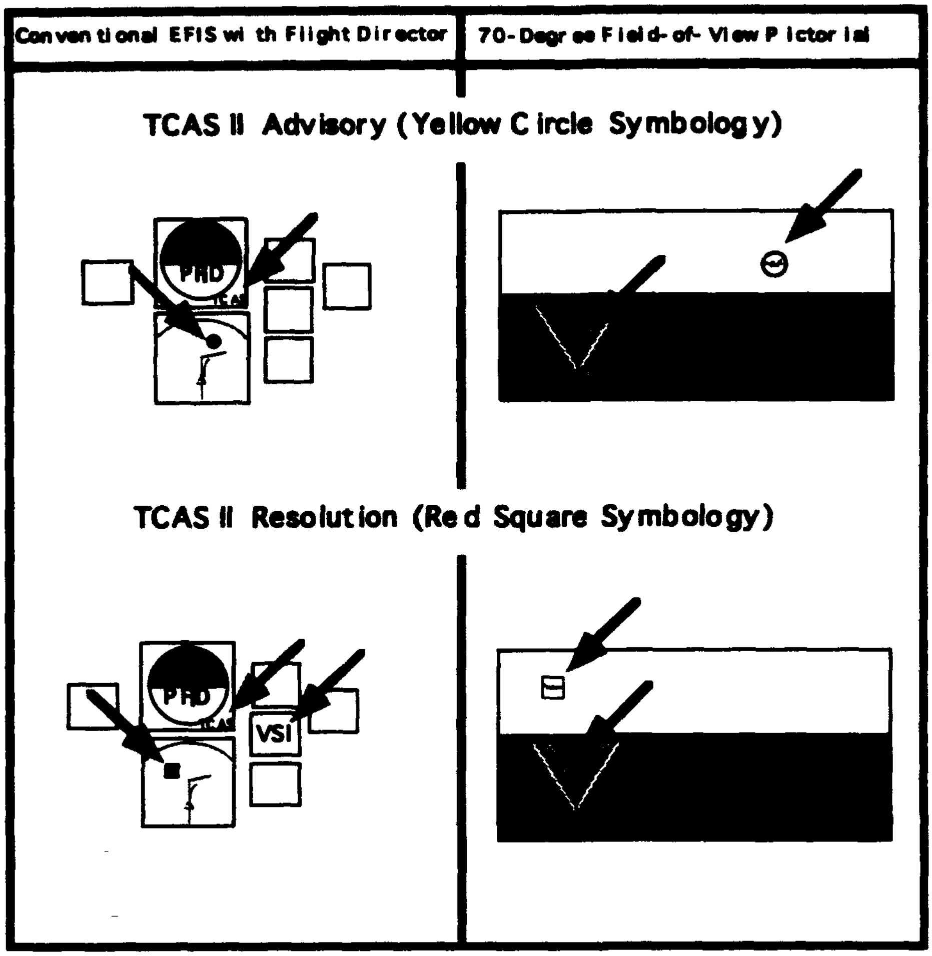


Figure 5. TCAS II Advisory and Resolution displays.

In order to evaluate spatial awareness, scenarios (to be discussed in the next section) were constructed that required the use of Traffic Collision and Avoidance System (TCAS) II implementations. Therefore, both display types (all four conditions) incorporated TCAS symbology, but the implementation differed with respect to the TCAS command portion (Figure 5). The conventional displays incorporated TCAS symbology on the ND, along with relative altitude tags and vertical direction (if climbing or descending).

The meaning of the symbology itself is defined, for purposes of this experiment, in Table 1. In actual field service, the TCAS advisory algorithms have changed and their implementation have become more sophisticated since the inception of this experiment. For the conventional displays, the TCAS command to either Climb or Descend was implemented on the Vertical Speed Indicator (VSI) in the form of a color-coded command bar. The pilot

responded by keeping the VSI needle in the green-colored portion of the indicator (and out of the red). When this was achieved, the pilot was following the TCAS command at an appropriate vertical rate. Warnings and commands were strictly visual (no auditory displays). For the pictorial displays, TCAS symbology was implemented in the same manner on the secondary see-through ND with one important augmentation. The computer-generated traffic in the OTW scene was also enclosed in a TCAS symbol with the appropriate warning color- and shape-coding. No resolution command (i.e., no vertical speed command) was presented with the pictorial formats. For all four display conditions, the TCAS was turned off below 500 feet AGL, although unfilled blue square symbols were used to represent other traffic on the displays, and their positional information continued to update.

Situation Awareness Assessment Tools And Techniques

The assessment of "situation awareness" is probably much more difficult than any attempted definition. Several techniques have been suggested in the overall literature, each with its own advantages and drawbacks. The most common method is by measuring traditional pilot/vehicle performance. However, there has been no established *direct* relationship between performance and awareness and, therefore, these measures should be supplemented by additional techniques (Sarter & Woods). The following is a list of additional techniques, compiled from Tenney, Adams, Pew, Huggins, & Rogers (1992), that were considered.

"Think-Aloud" Protocols

With this technique, subjects are encouraged to verbalize what they are thinking and describe what they are doing and why. It is considered a somewhat intrusive technique and is utilized only if the subject tends to do this anyway. The experimenter takes notes and compares what the subject says to what the subject does.

Anomalous Cues/Detection Time

This technique requires setting up scenarios that introduce slowly developing problems that may require some subject interaction. The experimenter then measures the time it takes for the subject to detect the problem, as well as the time before any corrective action is taken.

Freezing/Probes

This method entails a direct approach in which the experimenter either interrupts a task or "freezes" the task and then proceeds to take some form of measurement. Usually, the experimenter asks the subject relevant questions (in effect, probing them) concerning the task the subject was performing (Endsley, 1988; Endsley, 1989). Often questions are asked as to future events (based on what has transpired until the moment of "task freezing"), which may provide greater insight as to the subject's awareness of the situation at that moment. In other

words, the better the SA, the more accurately the subject will be able to predict the immediate future. In addition, after resuming the task, other measurements indicative of SA may be taken (such as time to restore to some predetermined condition). These methods require caution in that not only has the original task been corrupted, but the probe results must rely on the subject's short term memory.

Static Image Flash/Quiz

This method, simply stated, involves subject recognition of static information, scenarios or conditions when presented over a short period of time. The supposition is that the more accurately the subject is able to perceive or recognize the situation thus presented, the better the SA as provided by that particular information display system.

"Garden Path"/Detection Time

This technique involves leading the subject to an erroneous conclusion (by slowly developing parallel events) and then measuring the time it takes for the subject to detect the mistake in interpretation (i.e., the subject is presented information in such a way that a failure is correctly realized; however, it is attributed to the wrong source). Scenarios for this technique are more difficult to formulate.

Subjective Methods

Subjective methods mainly consist of questionnaire type evaluations where the subject, either verbally or by handwritten means, expresses personal opinions or feelings about the topic.

Techniques Selected

For this experiment, several techniques from the above list were chosen based upon the ability to generate of suitable scenarios in the context of transport approach and landing operations. The traditional lateral and vertical RMS errors, as well as control input data, were recorded directly during the basic or standard task, which was flying a Standard Terminal Arrival Route (STAR). All of the scenario tasks were implemented within the standard task. Two conflicting traffic scenarios were generated in order to utilize the anomalous cues/detection time technique. These scenarios included crossing traffic situations that caused TCAS alerts, as well as runway blunders by traffic on landing approach to a parallel runway. Two types of probe techniques were also employed. The first technique (which enabled two scenarios) interrupted the standard task by blanking the displays and then introduced a new task – flying with a backup instrument (only the eight-ball portion from the EFIS primary flight display). The supposition to be tested was that a superior display format would allow the pilot to think ahead of the airplane and thus be able to continue flying based on retained information. This scenario was thus formulated as a "Think-Ahead" awareness tool. The Blanking Scenario was followed immediately by the Offset Scenario, in which the aircraft was offset to one of four predetermined positions around the planned flight path and

then the time to restore to the intended flight path was measured. The Offset Scenario can be classified as a "Task Interruption/Time-to-Restore" scenario. The other probe technique utilized a new, unanticipated task, (a new STAR) that was frozen during execution, and relevant questions about spatial orientation, instrument readings and traffic awareness were then verbally presented. Finally, numerous subjective questionnaires were administered in which the subject evaluated the displays subjectively by answering relevant questions and by ranking the displays based upon the perception of the awareness afforded. Unsolicited subject comments were also recorded throughout the trials. Further explanations of the individual scenarios, SA evaluation techniques and measures follow in the next section on experimental tasks.

Naive Versus Experienced Replications

In utilizing the "Anomalous Cue/Detection Time" technique or some of the probe techniques, a major concern arises in the need for repetitions of the technique across the experimental conditions of interest, and replications within each of those conditions for statistical purposes. The problem is one of expectation by the subject of the occurrence of such an incident. These techniques are probably most effective when the subject is taken completely unaware (i.e., he is "naive"). However, this approach needs many subjects for the required between-subjects experimental design, especially when multiple factors are involved. In this effort to measure spatial awareness, an approach was taken in which some of the desired comparisons could be made using naive data, but comparisons of naive versus experienced results would also be available to judge the magnitude of the effect of prior exposure to the technique. Then most of the comparisons of interest would be made with a more economical within-subjects experimental design utilizing only the experienced data.

Experimental Tasks And Schedule

Sixteen pilots, all with extensive glass-cockpit experience, and most of whom were current line pilots with national commercial airlines (three were test pilots with commercial airplane manufacturers), acted as subjects in the experiment. Six separate experimental tasks based on the selected SA assessment methods, as previously discussed – actually six separate experiments – were embedded within the spatial awareness assessment efforts. These tasks included the Standard Approach Task, the Traffic Conflict Scenario, the Runway Blunder Scenario, the Blanking Scenario, the Offset Scenario, and the Probe Approach Task.

Standard Approach Task

All of the scenario tasks mentioned above were implemented within the Standard Approach Task. This basic task was a simulated STAR about 27 nautical miles (nm) in length, consisting of a complex, MLS-type approach (Figure 6) to closely-spaced, parallel runways.

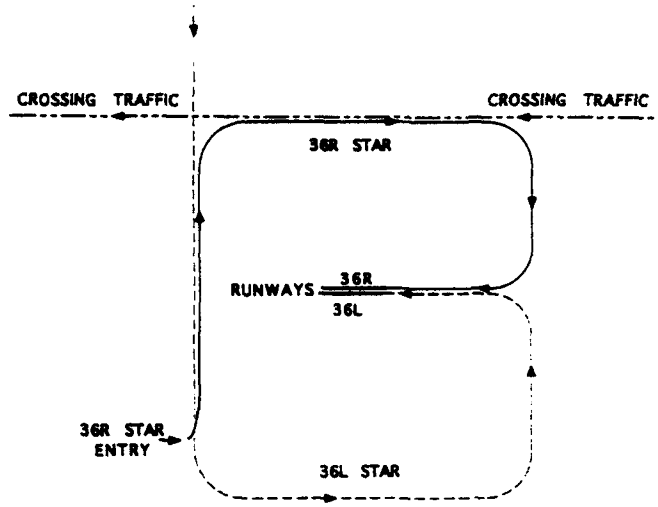


Figure 6. MLS-type standard approach to right runway.

The short final approach segment was only 1.7 nm in length. The STAR, the neighboring traffic routes (Figure 7), and the runway configuration (Figure 8) were constructed to provide a very complex environment of sufficient duration (about ten minutes per flight) for exercising the selected SA measurement tools. The environment was not intended to replicate the real world, but merely to represent a somewhat realistic, demanding future environment.

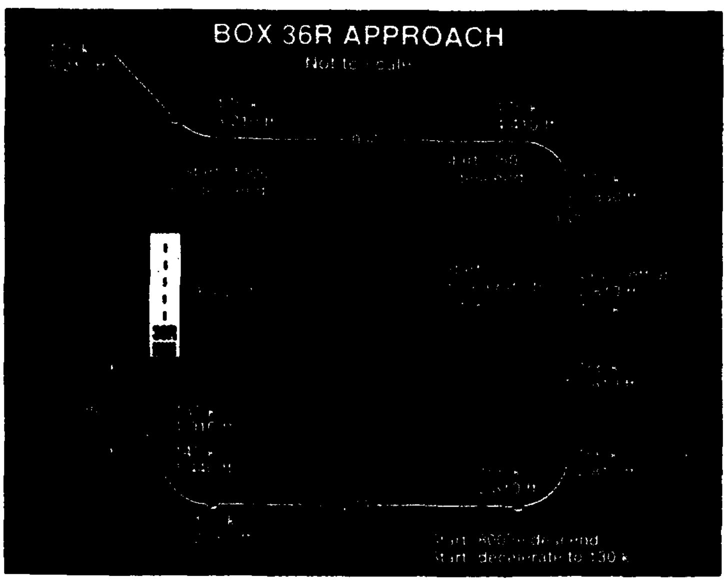


Figure 7. Traffic routes of parallel approach and crossing traffic.

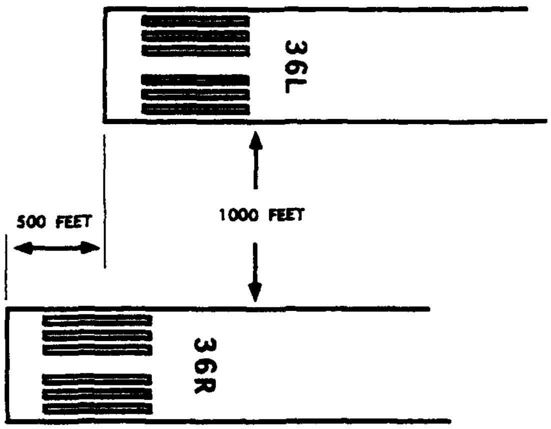


Figure 8. Offset, parallel runway configuration.

Active traffic was included on all routes, with several aircraft both preceding and following the ownship on the basic STAR, and with a constant stream of traffic on the STAR leading to the parallel runway, as well as with occasional traffic on the crossing route.

The pilot's task was to fly the STAR manually (including throttle inputs) using the head-down display condition available. While it is recognized that conventional EFIS displays are not used to fly below decision height altitudes in the real world (e.g., 200 feet) without an out-the-window transition, for the purposes of this investigation, the flight ended at the threshold without any transition. All of the awareness scenarios involved in the investigation were completed well before a 200 foot altitude was obtained. The STAR was divided into segments for analysis purposes (Figure 9), and the performance metrics for the standard task were the traditional lateral and vertical path tracking performances. While these measures are not really spatial awareness measurements, they are of related interest.

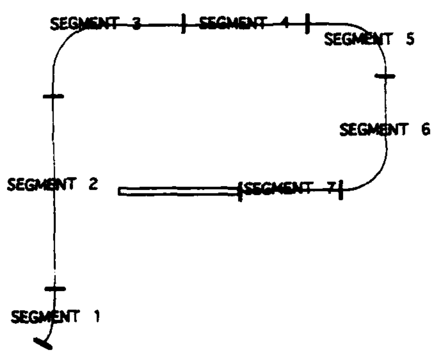


Figure 9. Segmentation of the Standard Approach Route for statistical analyses purposes.

The Traffic Conflict Scenario

The basic approach pattern (to the right parallel runway) always included other aircraft flying a STAR to the left runway. For the Traffic Conflict Scenario, which each pilot encountered in the data collection session only once for each Display Condition, any one of two aircraft, flying an opposing heading from ownship on segment 2 of the basic STAR (Figure 9), would inexplicably initiate an altitude maneuver which was intended to lead to a TCAS advisory situation for the ownship. The performance metrics for this scenario were the detection time (from the beginning of the altitude maneuver by the approaching traffic to the pilot's announced detection of the resulting threatening situation) and maneuver time (from the beginning of the altitude maneuver to the initiation of an avoidance maneuver, if one was initiated, by the ownship pilot). A data run was ended after the completion of the Traffic Scenario without continuing to the threshold. This scenario can be classified as an "Anomalous Cue/Detection Time" scenario in which the effect of prior exposure can be significant. Therefore, naive data was surreptitiously collected during the training sessions, with the pilots blocked across the Display Conditions (four pilots first encountered the scenario with each specific display). Later, the pilots became well-trained for the scenario.

The Runway Blunder Scenario

The basic approach pattern always included another aircraft landing on the left runway thirty seconds ahead of ownship (landing on the right parallel runway). For the Runway Blunder Scenario, which each pilot encountered in the data collection session only once for each Display Condition, that lead aircraft would inexplicably leave his landing pattern and cross in front of the ownship's flight path during final approach (while the ownship's planned altitude was 400 feet AGL. The TCAS advisory and resolution logic was turned off below 500 feet, although the appropriate displays still presented the other traffic with the unfilled blue square symbology). The performance metrics for this scenario were the detection time (from the beginning of the crossing maneuver by the neighboring traffic to the pilot's announced detection of the resulting threatening situation) and the maneuver time (from the beginning of the crossing maneuver to the initiation of an avoidance maneuver, if one was initiated, by the ownship pilot). This scenario can also be classified as an "Anomalous Cue/Detection Time" scenario, and surreptitious data was collected to allow naive versus experienced contrasts to be examined. RMS tracking data collection was ended before initiation of the Runway Blunder Scenario.

The Blanking Scenario

The Blanking Scenario exposed each of the sixteen pilots to four incidents of simulated display system failure for each Display Condition. The pilot's sole source of information with which to continue flying the ownship in these cases was a backup instrument (the eight-ball portion). Tracking data was collected for fifteen seconds of backup instrument flying and the performance metrics for this scenario were the RMS vertical and lateral tracking errors during that fifteen seconds. The Blanking Scenario, activated in segment 3 of Figure 9, was followed immediately by the Offset Scenario (activated in segment 4). The standard RMS tracking performance measures were not gathered for segments 3 and 4 during an approach

that included the Blanking and Offset Scenarios. However, tracking data collection was resumed after path recovery for the remaining segments of the flight (segments 5 - 7).

The Offset Scenario

The Offset Scenario exposed each of the sixteen pilots to four incidents of simulated recovery from display system failure for each Display Condition. After flying the backup instrument through the Blanking Scenario, the instrument screen would go totally blank for ten seconds, after which the original Display Condition would reappear. Upon reappearance, the position of ownship relative to the desired flight path was totally independent of the flying performance obtained with the backup instrument. The pilot's task in this scenario was to determine where the ownship was relative to the desired flight path, and to then return to the flight path in a timely manner, remembering that the vehicle simulated was a passenger airliner. The performance measure for this scenario was recovery time (with a return to path defined as achievement of less than half a dot error in lateral and vertical tracking and a heading error of less than 5 degrees).

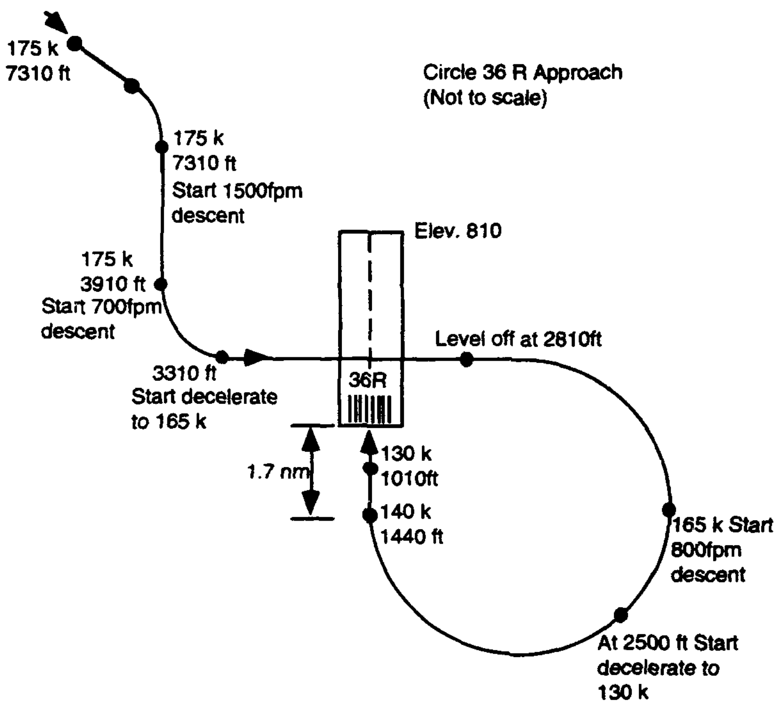


Figure 10. Non-standard approach utilized for probe technique analysis.

The Probe Approach Task

The probe technique of introducing a new task to the subject and then freezing the task during execution and conducting an extensive quiz to assess awareness was implemented in this study by using a new STAR. Three probe STAR's were flown by each pilot, using two of the display conditions. The STAR represented in Figure 10 was used as the "naive" probe at the end of the training session for each pilot.

The same STAR, but with a runway elevation offset of 710 feet (which changed all of the waypoint altitudes), was used as the "experienced" probe for the same Display Condition at the end of the data collection session. The pilots were blocked across the Display Conditions (four pilots first encountered the naive probe with each specific display). Another STAR (Figure 11, a mirror image of the original STAR) was also used in the middle of the data session as an "experienced" probe with another Display Condition. Because of the extensive time consumed by the total experiment, only the display comparisons listed in Table 2 were provided by the Probe Approach Task.

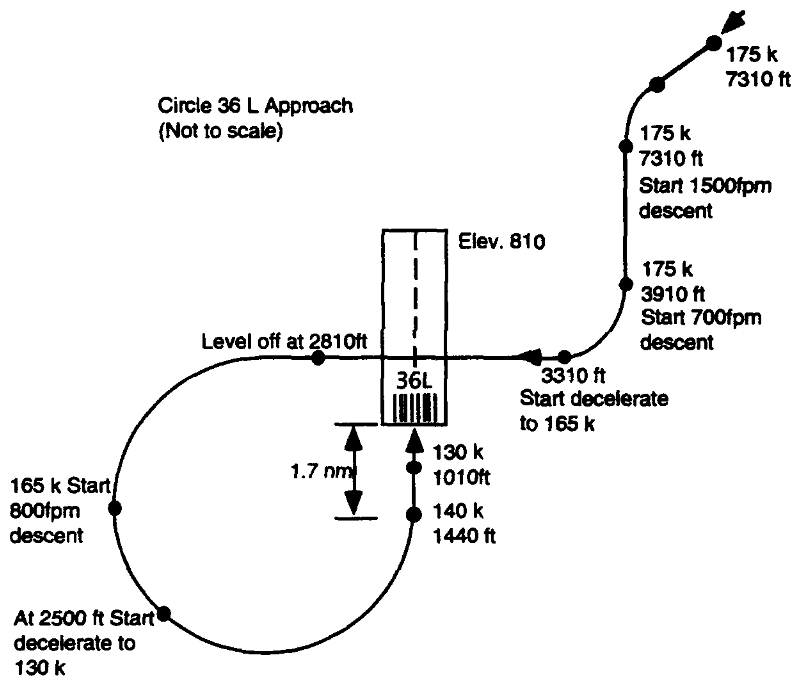


Figure 11. Non-standard approach utilized as second replicate for probe technique analysis.

Schedule

Table 3 presents the two day schedule of the experiment for an individual pilot. After being briefed on the purpose of the experiment, the details of each Display Condition, and the

familiarize himself with the handling characteristics of the airplane model in unstructured flight maneuvers. Half of the pilots used the conventional EFIS without Flight Director Display Condition for this purpose, while the other half used the 70 degree pictorial Display Condition. The pilots were then thoroughly trained with the standard approach task, and then were thoroughly exposed to each Scenario Condition, for each Display Condition. Surreptitious naive data was collected for the initial exposures to the Traffic Scenario and the Runway Blunder Scenario. At the end of the training session, the Naive Probe was administered.

The second day was the data collection session. The Display Conditions were randomly blocked across pilots, and the experimental tasks were randomized within each Display Condition. Table 4 presents an outline of a typical session, the details of which varied from pilot to pilot.

Table 2

- Probe STARs
  - 3 Landing Approaches per Pilot
    - 1 Naive Exposure
    - 1 Direct Comparison (Same Display)
    - 1 Direct Comparison (Different Display)

Probe Contrasts

- Naive vs. Experienced
- Display Comparisons
  - Conventional
  - With vs. Without Flight Director
  - With Flight Director vs. 70° Pictorial
  - Without Flight Director vs. 70° Pictorial

Table 3

Day 1 (approximately 10 hours)

- Briefing Session
- Training Session
  - Handling Characteristics Familiarization
  - Display Condition #1-4
  - 2 Surreptitious Naive Data Runs (TAM, Builder)
  - 1 Naive Probe

Day 2 (with rest periods, approximately 10 hours)

- Data Collection Session
  - Display Conditions #1-4
  - Questionnaires

Table 4

DISPLAY CONDITION	APPROACH CONDITION	QUESTIONAIRES			
		INTRUSION	DISPLAY EVALUATION	PROBE	DISPLAY COMPARISON
# 3	$R^*,O_4,O_1,R,O_3,T,R,O_2$	✓	✓		
# 1	$R,O_1,T,R,O_3,O_4,R^*,O_2,P_2$	✓	✓	✓	
# 2	$R,O_3,T,R,O_2,R^*,O_4,O_1$	✓	✓		
# 4	$R,O_2,R^*,O_4,O_1,R,O_3,T,P_3$	✓	✓	✓	✓

R = STANDARD APPROACH

R* = STANDARD APPROACH WITH RUNWAY BLUNDER

T = TRAFFIC AVOIDANCE MANEUVER

O_x = OFFSET OCCURRENCE

P_2 = PROBE RUN # 2 - Aware, Different Display Condition

P_3 = PROBE RUN # 3 - Aware, Original Display Condition

Experimental Results And Discussion

Most of the scenarios under investigation were designed as a full-factorial, within-subjects experiments, with Pilots, Display Condition, any Scenario Conditions, and Replicates as the factors. The data collected in the experiments were analyzed using univariate analyses of variance for each metric. The more important objective results are presented and discussed for each scenario, and some of the subjective results are discussed thereafter.

Traffic Scenario

Objective Results. Figures 12 and 13 present the results of the Traffic Scenario graphically. All sixteen pilots detected each threatening situation, regardless of the Display Condition (Figure 12). However, the differences between the detection times for the EFIS Display Conditions and the pictorial Display Conditions (Figure 13, about 10 seconds) were statistically significant. Differences within the display types (EFIS and pictorial) were not significant.

The altitude maneuver executed by the approaching traffic usually resulted in a TCAS advisory or a TCAS resolution, with the outcome dependent upon the current tracking performance of the ownship. There were thirty-eight cases in which an ownship avoidance maneuver was executed (Figure 12) and twenty-six cases in which the pilot decided not to execute a maneuver. The maneuvers may have resulted from a TCAS resolution or from an independent decision of the ownship pilot. The no-maneuver decisions may have been made because the situation was judged to be not serious. Detailed analysis of the data to determine the TCAS condition for each of the sixty-four cases has not yet been accomplished. In any case, the analysis of variance for the maneuver time measure found no statistically significant differences for any of the factors of the scenario experiment.

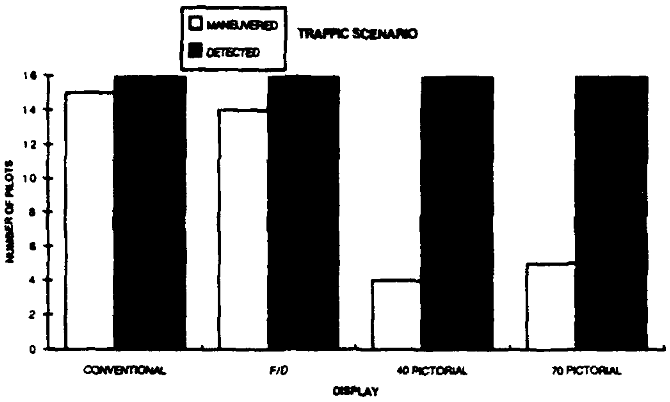


Figure 12. Traffic scenario number of detections and subsequent maneuvers by pilot.

Discussion of Objective Results. One may infer from these results that the pictorial displays provided the pilot with better traffic information than the EFIS displays. Detection of the threatening traffic situations occurred earlier and at greater distances (the 10 second earlier detection time translates into 1.0 nautical mile of increased separation) with the pictorial displays; and with the increased awareness of the situation, fewer avoidance maneuvers were initiated by the pilots.

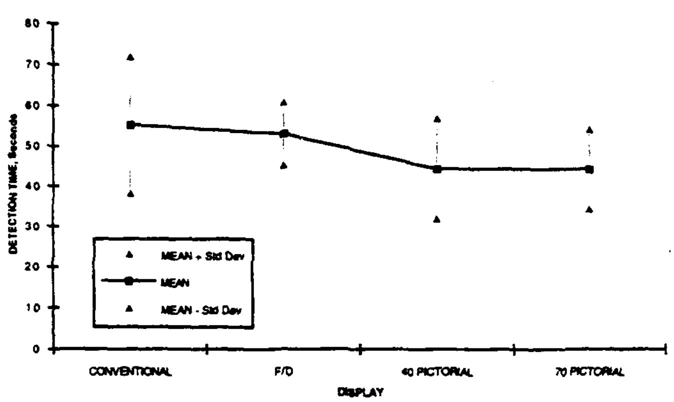


Figure 13. Traffic scenario mean detection times per display concept/condition.

Runway Blunder Scenario

Objective Results. Figure 14 illustrates the Runway Blunder Scenario and the obvious visual advantages of the pictorial display formats. With the pictorial displays, all sixteen

pilots detected each threatening situation (Figure 15). With the EFIS displays, only about half of the blunders were detected. The differences between the mean detection times for the EFIS Display Conditions and the pictorial Display Conditions (Figure 16, about 8 seconds) were statistically significant. The difference within the EFIS display types (3.3 seconds sooner for the Flight Director condition versus the without Flight Director condition) was also significant, while the difference between the pictorial conditions (0.7 seconds sooner for the 40 degree condition) was not significant.

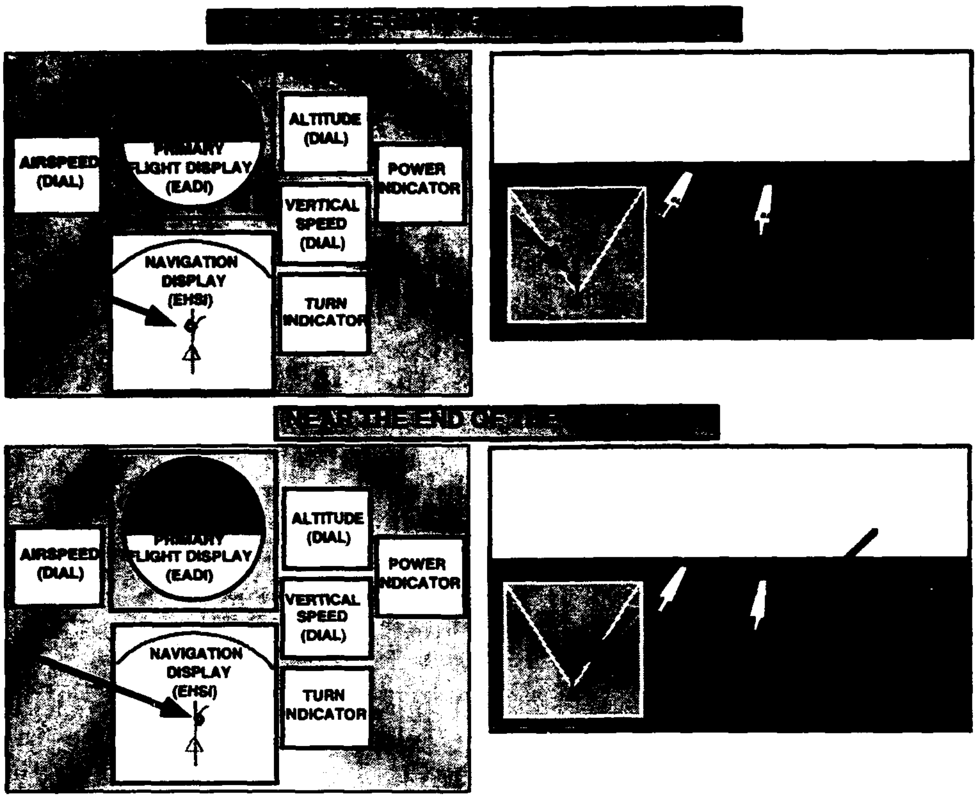


Figure 14. Runway Blunder scenario depicting parallel traffic incursions. In the conventional display format the incurring traffic is rendered as an unfilled blue diamond. In the pictorial display format it is rendered as an aircraft silhouette enclosed by a black square.

Of the sixty-four incidents of runway blunder, fifteen went undetected (all under EFIS-type Display Conditions, Figure 15. Within the forty-nine detected incidents, the pilots chose to initiate a go-around maneuver in thirty-one cases. Analysis of the maneuver time measure for those thirty-one cases revealed significant differences between most paired means comparisons (Figure 17). The maneuver time difference between the EFIS-type displays was statistically significant, with the Flight Director mean being 4.6 seconds earlier than the EFIS without Flight Director mean. The 40 degree pictorial display mean was a significant 3.4

seconds earlier that the Flight Director mean. The difference between the 40 degree and the 70 degree pictorial display means was not statistically significant.

Discussion of Objective Results. One may infer from these results that the pictorial displays provided the pilot with better traffic awareness nearing the runway than did the EFIS displays. Fifteen of the thirty-two incidents of the runway blunder went undetected with the EFIS displays, and when detection did occur, it came later than with the pictorial displays. And with the increased awareness of the runway situation, a lesser percentage of go-round maneuvers were initiated by the pilots when utilizing the pictorial displays.

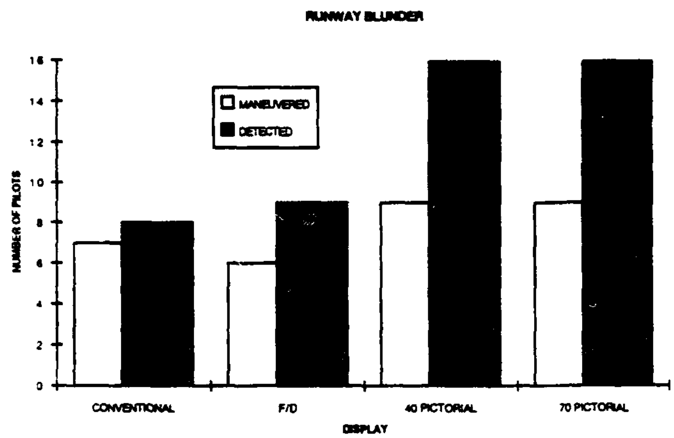


Figure 15. Runway Blunder scenario number of detections and subsequent maneuvers by pilot.

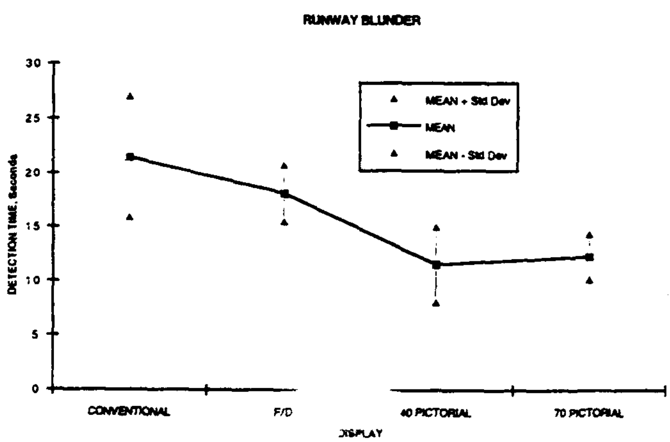


Figure 16. Runway Blunder scenario mean detection times per display concept/condition.

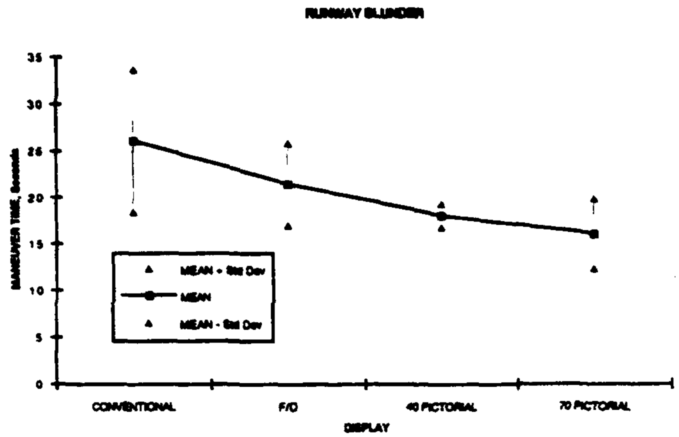


Figure 17. Runway Blunder scenario mean times to maneuver since detection.

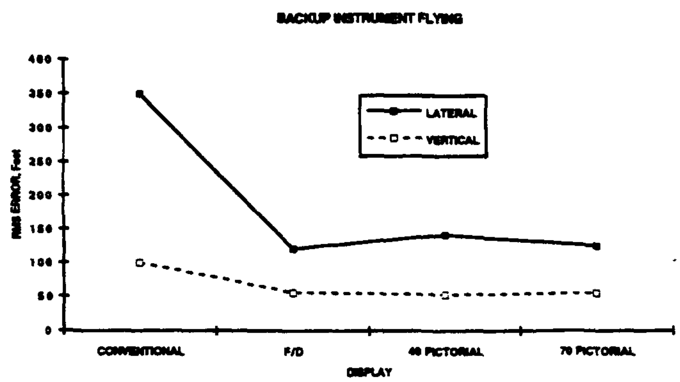


Figure 18. Blanking Scenario lateral and vertical mean errors (RMS) per display concept/condition.

Blanking Scenario

Objective Results. The Blanking Scenario results are presented in Figure 18 for the Display factor for the two performance measures, which both indicate poorer performance when flying the backup instrument after having reverted to that instrument from the conventional EFIS without Flight Director display.

Discussion of Objective Results.- One may infer from an in-depth analysis of these results that the Blanking Scenario yielded no meaningful results in terms of spatial awareness differences. It is much more likely that the differences presented in Figure 18 are caused by the large variations in the initial conditions at the time of reversion to backup instrument flying than by spatial awareness effects. For example, flying only lateral and vertical raw error information through a turn rather than following the flight director commands through the turn would result in vastly different variations in starting points for backup instrument flying just prior to turn exit. Therefore the lack of control of initial conditions is aliased with the other experimental factors, and the Blanking Scenario has yielded no insight as to spatial awareness differences among the various display concepts.

Offset Scenario

Objective Results. For the Offset Scenario (Figure 19), more time is required to recover when flying the conventional EFIS displays without Flight Director. With Flight Director, the recovery time was a significant 14.6 seconds quicker than the without Flight Director case, and the performances with the pictorial displays were at least 10.2 seconds faster than the Flight Director results (statistically significant). The difference between the pictorial conditions (2.5 seconds faster for the 40 degree condition) was not significant.

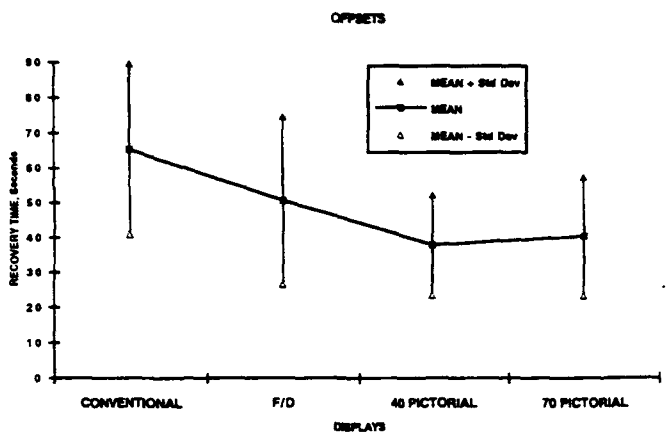


Figure 19. Offset Scenario mean times to recover to the intended flight path per display concept/condition.

Discussion of Objective Results. One may infer from these results that the pilots were able to determine where the ownship was located relative to the desired flight path and then returned to the flight path more quickly with the pictorial displays. The Flight Director recovery was faster than the conventional without Flight Director display, probably because interpreting the raw error information was more time consuming than just following the Flight Director commands. The difference in recovery time between the pictorial displays

and the Flight Director display could also be attributed to more aggressive manual intercepts of the flight path with the pictorial displays versus the intercept logic within the Flight Director.

Subjective Results

Obviously, with eleven questionnaires composed of numerous questions each, only a summary of the subjective results is possible for the purposes of this paper. Figure 20 presents rating results for two subjective categories, as typical examples. The pilots were asked to rate, on a scale of from very hard to very easy, the ease of becoming disoriented, and, in an opposite connotation (as a sanity check, the same question), the ease of maintaining spatial awareness, when using each display configuration (without comparison to the other display configurations). A dramatic improvement in both instances is provided by the two pictorial formats, and in particular, by the large-screen 70 degree version. Figure 21 presents the results of comparative rank ordering by the pilots for several categories, on a scale of from 1 (being the most desirable display, to 10 (being the least desirable display). The mean ranking is presented, along with the maximum and minimum rankings (not plus or minus the standard deviations). The categories presented compare the display concepts over all scenarios of the experiment, and include effectiveness in monitoring traffic, in reducing pilot workload, and the overall ranking for the entire experiment. Again, based on pilots' subjective data, a substantial improvement in all aspects of spatial awareness was provided by the two pictorial formats (in both mean ranking and spread), and in particular, by the large-screen 70 degree version.

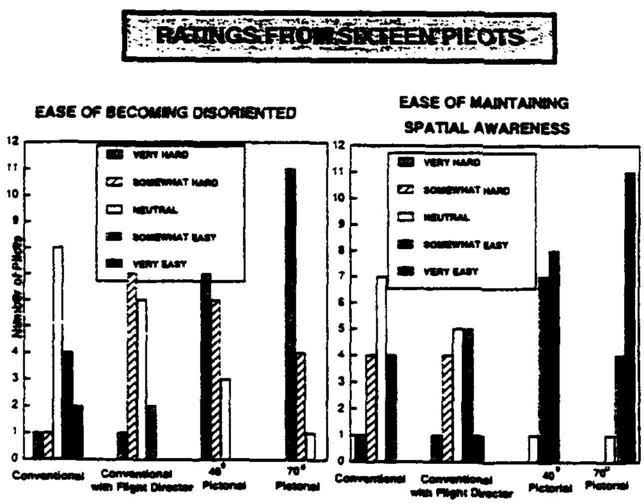


Figure 20. Rating results from the sixteen pilots for two example subjective categories.

In addition to the formal questionnaire results, other subjective comments were obtained. Some of those that stand out (regarding the pictorial displays) are:

- "Like flying on a beautiful VFR day."

- "Provides immediate assessment of the situation..."
- "Ability to fly complex approaches is greatly improved."
- "Easier to detect traffic incursions and runway blunders."
- "Display of pictorial world is natural and easy to interpret."

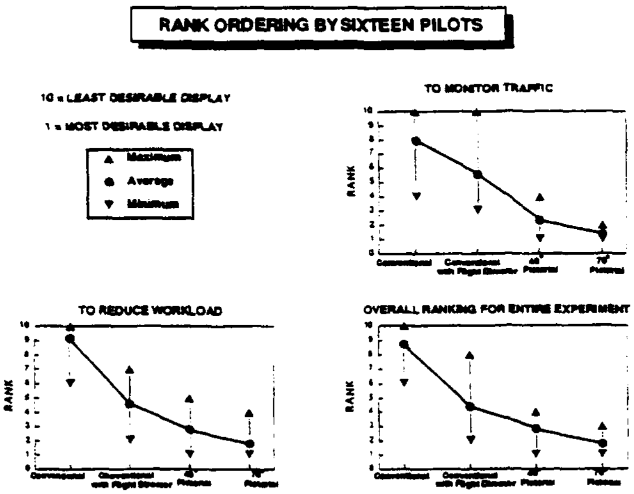


Figure 21. Comparative rank ordering by the sixteen pilots for three example categories.

Objective Tracking Performance Results

Objective Results. In addition to the more esoteric SA measurement techniques, standard RMS tracking data was collected and analyzed. The standard or basic STAR was divided into segments for this purpose (Figure 9), with the following designations:

SEGMENT DESCRIPTION
1. Turning Entry
2. Straight Descent
3. Descending Turn
4. Transition to Straight & Level, Decelerate
5. Level Turn
6. 3 degree Descent, Decelerate, & Turn Final
7. Short Final Approach

The analysis for Segment 1, the entry to the STAR from the off-path initial conditions, was not meaningful in terms of spatial awareness results, and is not presented. Figure 22 presents a comparison of the Display Condition RMS lateral error means from the sixteen pilots for each segment. Not surprisingly, lateral tracking performance error was significantly larger for every segment of the STAR for the conventional EFIS without Flight Director

display. Differences between the conventional EFIS with Flight Director and the two pictorial Display Conditions were also significant (Flight Director error was larger), while differences between the 40 degree and the 70 degree Display Conditions were not significant. Differences in performance between segments for a particular Display Condition can be attributed to the type of segment (segments 3, 5, and 6 included turns while segments 2, 4, and 7 were straight-aways).

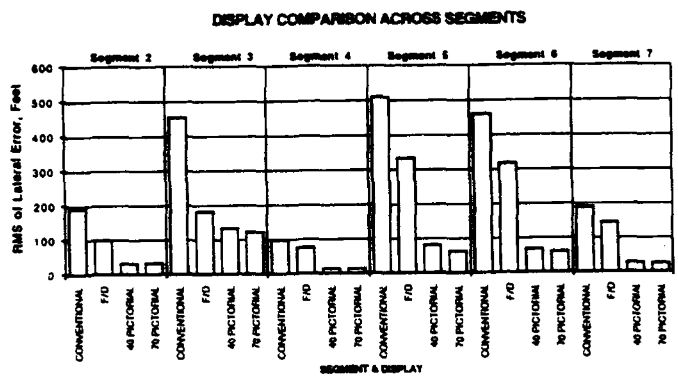


Figure 22. Mean lateral errors (for all sixteen pilots) per display concept/condition for each path segment.

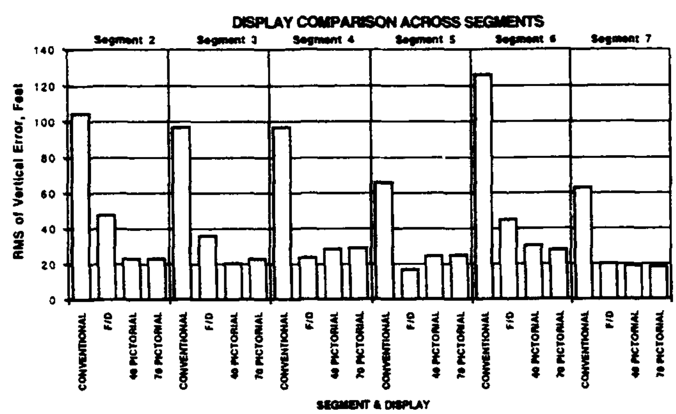


Figure 23. Mean vertical errors (for all sixteen pilots) per display concept/condition for each path segment.

Figure 23 presents a comparison of the Display Condition RMS vertical error means from the sixteen pilots for each segment. Not surprisingly, vertical tracking performance error was significantly larger for every segment for the conventional EFIS without Flight Director

Display Condition. Differences between the conventional EFIS with Flight Director and the two pictorial Display Conditions were also significant, but only for segments -,3, and 6 (Flight Director error was larger). Differences between the 40 degree and the 70 degree Display Conditions were not significant for any of the segments. Differences in performance between segments for a particular Display Condition can be attributed to the type of segment (segments 2,3, 6, and 7 included descents while segments 4 and 5 were level).

Discussion of Objective Results. The lateral and vertical tracking performances with the pictorial Display Conditions were at least as good or better than the performances with the EFIS Flight Director Display Condition, and the Flight Director performances were much better than the EFIS without Flight Director condition. While these facts lead to no conclusions about increased or decreased spatial awareness, they do provide the assurance that the increased spatial awareness provided by the pictorial displays, as measured by the other measurement tools, was not gained at the expense of degraded tracking performance.

Inferences From Results

Conclusive inferences can be drawn from the objective and subjective results available at this time for comparisons between the two EFIS display formats, between the two pictorial display formats, and between the conventional EFIS displays and the pictorial displays.

EFIS Comparisons

In all cases in which objective or subjective results compared the EFIS displays with and without Flight Director, either equivalent or better performance was achieved with the Flight Director EFIS display. Better spatial awareness appears to be gained through the lower path-tracking workload imposed by the Flight Director, which allows time for scanning sources of information other than the Flight Director needles. Flying raw data error in the EFIS without Flight Director condition requires that almost constant attention be devoted to the path-tracking task.

Pictorial Comparisons

The objective data revealed equivalent or slightly better performance for the 70 degree pictorial display compared to the 40 degree FOV. The subjective data revealed a stronger preference for the wider FOV, particularly for awareness during turn entry and traffic situations.

EFIS and Pictorial Comparisons

Both the objective and subjective data demonstrated that the integrated pictorial displays provided increased spatial awareness over the conventional EFIS display formats.

Concluding Remarks

A simulation study was conducted using sixteen commercial airline pilots repeatedly flying complex MLS-type approaches to closely-spaced parallel runways to compare the spatial awareness of pilots flying with conventional flight displays to their awareness when flying advanced pictorial, "pathway-in-the-sky" displays. Various situational awareness measurement techniques (which were incorporated within scenarios), involving conflicting traffic situation assessments, main display failures, and recoveries from unknown positions, were used to assess the pilots' spatial awareness with the different display formats, both objectively and subjectively. The numerous spatial awareness tools utilized in the experiment proved to be most effective in the assessment (with the exception of the Blanking Scenario), in that the results were consistent across and within the objective and subjective measures.

Analyses of the data for the Traffic Scenario and the Runway Blunder Scenario involving the surreptitious naive data runs mentioned earlier have not yet been completed. These analyses will examine display comparisons upon naive exposure to the scenarios utilizing between-subjects experimental designs. Naive versus experienced contrasts across Display Conditions will also be examined. Data from the Probe STAR's, which will also examine naive versus experienced performances, as well as provide display comparisons, have not been analyzed yet, either.

Nevertheless, the objective data analyses thus completed revealed that better spatial awareness performance was usually achieved with the Flight Director EFIS display compared to the without Flight Director EFIS display. However, the major objective results of the study were that the integrated pictorial displays consistently provided substantially increased spatial awareness over either of the conventional EFIS display formats. The wider FOV pictorial display gave slightly better objective results than the narrower pictorial format.

A summary of the numerous subjective results would indicate a very strong preference for the Flight Director presence within the EFIS displays. But again, the major results of the study were that a dramatic improvement in all aspects of spatial awareness is provided by the two pictorial formats, and in particular, by the large-screen 70 degree version.

Integrated pictorial displays have therefore shown significant promise for providing improved situation awareness and corresponding safety benefits. These types of formats are expected to provide the cornerstone for an effective synthetic vision system, a system which is an enabling technology for solving restricted visibility problems associated with advanced subsonic and future high speed civil transports.

References

Abbott, T. S. (1989). Task-Oriented Display Design: Concept and Example. *SAE Technical Paper Series 892230, Aerospace Technology Conference and Exposition.* Anaheim, CA, September 25-28.

Brahney, H. (1992). Aerospace technology: to the 21st Century. *Aerospace Engineering (An SAE Publication)*, Pages 16 -17.

Dorighi, N.S., Grunwald, A. J. & Ellis, S. R. (1992). Perspective Format for a Primary Flight Display and its Effect on Pilot Spatial Awareness. *Proceedings of the 11th Digital Avionics Systems Conference, Crew Station Technology Session,* Seattle, WA.

Dorighi, N. S. & Ellis, S. R. (1991). Evaluation of Perspective Displays on Pilot Spatial Awareness in Low Visibility Curved Approaches. *AIAA.*

Endsley, M. R. (1988). Situation Awareness Global Assessment Technique (SAGAT). *Paper presented at the National Aerospace and Electronic Conference (NAECON),* Dayton, OH.

Endsley, M.R. (1989). A Methodology for the Objective Measurement of Situation Awareness. *Paper presented at the AGARD Symposium on Situation Awareness in Aerospace Operations,* Copenhagen, Denmark.

Hatfield, J.J. Parish, R. V. (1990). Advanced Cockpit Technology for Future Civil Transport Aircraft. *Proceedings of the Eleventh Annual IEEE/AESS Dayton Chapter Symposium "The Cockpit of the 21st Century - Will High-Tech Payoff?" Advanced Displays and Controls Topic.*

Proceedings of the 7th Plenary Session of the Synthetic Vision Certification Issues Study Team. (1992). Williamsburg, VA, June 23-34.

Regal, D. M., Rogers, W. H. & Boucek, Jr., G. P. (1988). Situational Awareness in the Commercial Flight Deck - Definition, Measurement, and Enhancement. *7th Aerospace Behavioral Technology Conference and Exposition: SAE,* Warrendale, PA, Pages 65-69.

Regal, D. & Whittington, D. (1993). Synthetic Vision in Commercial Aviation - Display Requirements. *Proceedings of the Aviation Psychology Symposium,* Ohio State University, April 26 - 30.

Swink, J. R. & Goins, R. T. (1992). High-Speed Research System Study - Advanced Flight Deck Configuration Effects. *NASA Contractor Report 189650, Contract NAS1-19345,* June.

Sarter, N. B. & Woods, D. D. *Situation Awareness: Critical Phenomenon or Ill-Defined Concept?* Ohio State University, Cognitive Systems Engineering Laboratory

Tenney, Y. J., Adams, M. J., Pew, R. W., Huggins, W. F. & Rogers, W. H. (1992). A Principled Approach to the Measurement of Situation Awareness in Commercial Aviation. *NASA Contractor Report 4451, Contract NAS1-18788,* July, 1992.

Air Traffic Control

- Situational Awareness in Air Traffic Control
- Controlling Automation in Future Air Traffic Control: The Impact on Situational Awareness
- Mental Models and Situation Awareness in Air Traffic Control
- The Role of Dynamic Memory in Air Traffic Controllers' Situation Awareness
- SATORI: Situational Assessment Through The Re-Creation of Incidents
- The Impact of Associate Systems Technology on an Air Traffic Controller's Situational Awareness

Situational Awareness in Air Traffic Control

V. David Hopkin

Human Factors Consultant

Introduction

According to its Program, this conference follows a familiar pattern. It focuses on a single theme, in this case Situational Awareness, and views it in relation to a diversity of approaches, perspectives, contexts, and applications. The proceedings, published in their entirety or selectively, then represent the current state of knowledge of that theme.

Some aspects of this process inevitably may appear initially to be somewhat artificial and forced. In a few contexts, situational awareness may even convey the impression that it is a theme in search of applications. The idea that situational awareness is sufficiently universal as a concept to be relevant, valuable and applicable so widely seems rather unlikely. The idea that it is equally applicable so diversely seems even more unlikely.

For if it is so generally applicable, why is it so recent? It is not mentioned in Reber's (1985) Dictionary of Psychology. Yet reviews of situational awareness (e.g., Taylor, 1991; Garland et. al., 1991; Garland et. al., 1992) now collectively contain so many references on it that an article of the type published in the Annual Reviews of Psychology would be needed to mention them all, and could not treat them in depth.

A Unifying Concept

Speculations about the reasons for the sudden popularity of situational awareness start to provide an explanation of why it seems such a fertile notion and to justify a meeting devoted exclusively to it. Perhaps situational awareness is apparently such a recent concept because formerly it was described in other terms or assigned different names in different contexts. If this is accepted, the true potential value of situational awareness as a concept begins to emerge. It is a unifying concept, and perhaps a universal one, for it provides a means to reveal commonalties across contexts and applications, and to permit practices and constructs to be brought together and compared that previously were used and even named differently but that now seem to have much in common.

Situational awareness is also a unifying concept in another important respect: it treats as an indivisible whole a notion previously partitioned so that its various aspects were addressed

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

separately. These prospects for universality and unification offered by situational awareness fully justify the form which this conference takes and augur well for significant progress from the broad coverage provided by the proceedings. They show that my initial misgivings about the scope and rationale of the meeting were misplaced. It could be a landmark event.

Situation or Awareness

Many people have had the disconcerting experience, particularly while driving a car along a featureless but familiar road with little traffic, of suddenly becoming aware that they have no recollection of the last few miles they have driven. Does this experience constitute loss of situational awareness or does it not? The question is neither trivial nor trite. It points to an anomaly in the literature on situational awareness and in the nature and measurement of the phenomenon itself. Is the emphasis on the situation or on awareness of it?

If situational awareness primarily concerns the situation, then it is measurable in terms of performance, and the fact that any passenger in the car usually reports no noticeable change in driving performance during this gross lapse of attention suggests that situational awareness remains unimpaired because performance of the driving task remains intact. If situational awareness primarily concerns awareness, then in this driving condition situational awareness has been lost because awareness has been lost, and subjective measures are necessary to establish this. Different measures can give opposing evidence on when the phenomenon of situational awareness is present. In the literature, some of the claimed measures of situational awareness are of performance, and others are subjective. Trying to have this both ways engenders confusion about situational awareness itself, and potentially anomalous or contradictory findings about it. What sort of measures would be required to prove that two people had the same situational awareness?

In this instance of car driving (to pursue this example a little further), neither performance nor subjective measures would suffice. Performance measures are probably so insensitive that they cannot even distinguish reliably between the presence or absence of this gross lapse in awareness. Subjective measures are little better: it is necessary to rely on them for evidence of the lapse of attention but they are useless thereafter to dealing with a phenomenon the essence of which is that nothing occurring during it can be recalled. This relates to automaticity or automatization, which is featured in many theories or descriptions of situational awareness, sometimes as a claimed essential condition for situational awareness and sometimes as a kind of negation of situational awareness.

Too Much Situational Awareness

In considering situational awareness in air traffic control, perhaps the first issue to be addressed is whether the human factors work has been independent, or, having far fewer resources, has tagged along after the work on cockpits, and borrowed, adapted or appropriated studies of flight decks. This does not seem to have happened; rather, the work in air traffic

control has paralleled cockpit studies, using different terminology, notably the controller's picture. That explains why in this paper there is frequent cross-referral between general situational awareness issues and specific air traffic control topics, but not so much cross referral of either to cockpit environments.

Lapses of attention of the kind that occur during car driving are not commonly reported in air traffic control. This could be either because they are genuinely rare or because common prudence suggests that a controller might be unwise to publicize them. Air traffic control does however, pose a further question the answer to which seems to depend crucially on this apparent dichotomy between the situation and awareness of it. The question is: Can the air traffic controller have too much situational awareness?

While the preponderant view seems to be that it is not possible to have too much awareness, it is certainly possible to have too much situation that the controller must be aware of. This is the essence of the problem when the controller "loses the picture." This occurs when the amount of information exceeds the controller's capability to maintain it as an integrated and coherent entity for the purposes of the control of air traffic. It is associated with heavy traffic, when there is generally most information to be integrated. Its onset can be sudden, but is often preceded by a period during which the controller becomes increasing worried that loss of picture is liable to occur. If it does, the consequences operationally can be very serious because, once the picture has been lost, the controller can seldom recall it in its entirety again but has to rebuild it painstakingly aircraft by aircraft, often by systematically re-committing to memory the details of each aircraft under control, one at a time, knowing all the while that if this is successful not only will the picture have been restored but so will the self-same conditions which previously led to loss of picture.

Whether there can be too much situational awareness has profound consequences for measurement. If there cannot be too much situational awareness and it has no maximum then no valid measures can be devised that purport to deal with proportions or percentages of it. But if it has a maximum or optimum, such types of measure become feasible, at least in principle. In the former case it is never possible to prove that full situational awareness has been achieved. In the latter case, this may be possible.

"Unawareness"

At one time, the concept of 'unawareness' was used to describe a perceived human factors problem in aviation, and particularly in cockpits. The concept seemed to be most concerned with what would now be called 'attention', but the role of consciousness was dismissed as misleading, and the emphasis was on the measurement of behavior. An extract, converted to non-sexist language, may convey its flavor (Hopkin, 1967).

> A change in what a human is aware of may not change behavior; a change in behavior need not entail a change in conscious awareness. The most promising solutions to the problem of unawareness therefore depend on effecting a change in behavior, and this may or may not be associated with a change in conscious awareness.

Unawareness mistakes may occur for a number of reasons. The following categories may not be equivalent in terms of conscious awareness but are equivalent in terms of their consequences on human actions:

a. The human has a full appreciation of the situation but an inability to take action. This is rare but may occur, for example, in extreme fatigue where the human can appreciate a situation but be too tired to do anything about it.

b. The human may have an adequate perception of all the relevant stimuli but a failure to appreciate their meaning or import. The human may for example see a light but forget what it means.

c. A human may fail to perceive a particular stimulus. He/she may not notice a light for example, but see other things perfectly well.

d. A human may not perceive any of the surrounding stimuli, being for example, preoccupied with his/her thoughts and mind wandering."

The Controller's "Picture"

In air traffic control, situational awareness seems to correspond quite well with the concept of the controller's picture. This picture is sometimes construed as an example of a mental model but although the picture includes the controller's mental model it is not confined to it, being a more dynamic entity than most mental models in that it incorporates changing states and their consequences. In some respects, the "picture" seems specific to air traffic control. Typically the controller of a sector, that is a geographical region of airspace containing en route traffic at high flight levels, acquires the "full picture" over a period of about twenty minutes because by then the controller knows the full history of all the traffic while it has been under his or her direct control. However, an adequate picture for controlling the traffic can be built much more rapidly, and this occurs whenever one controller hands over the control responsibility for a sector to another controller at the end of the work shift, or whenever a supervisor adds an annotation to a controller's paper flight progress strip for a particular aircraft.

To anyone with no knowledge at all of air traffic control, the controller's workspace is meaningless and mystifying. This is because it contains no information about what it is, what it is for, or how and why it could or should be used. Its meaning depends therefore on what is known, as well as on what is portrayed. Only the controller can make sense of the portrayed information in terms of a "picture" of the air traffic. Therefore only a controller could have full situational awareness in an air traffic control workspace. The naive occupant of the workspace would see a room and furniture and seating and meaningless information displays and keys with unknown functions. Does this constitute a superficial form of situational awareness, or is situational awareness of an air traffic control workspace restricted in principle to those with professional knowledge of air traffic control? If the latter is required, according to what criteria could the trainee controller be said to possess sufficient knowledge

to have situational awareness, and what kind of knowledge would that be? The naive observer has a different kind of situational awareness or no situational awareness at all, and can have no insight into the level of complexity of awareness that the situation is capable of.

Individual Differences

Situational awareness therefore is learned, and there are gross individual differences in actual situational awareness and perhaps also in the potential for situational awareness in different individuals. Several consequences follow from this, which can be stated as hypotheses:

1. Different controllers, given the identical displayed information, will have different situational awareness so that the situational awareness of one controller might not include everything that another controller would assume it contained.

2. Because situational awareness depends on learning, it should to some extent be explicable and predictable in terms of learning concepts and theories, provided that they are valid.

3. Situational awareness, as the product of a learning process, can be built, extended, developed, entrenched and reinforced.

4. Situational awareness, also as the product of a learning process, can be fallible, incomplete, distorted, subject to error, or forgotten.

5. The capability to achieve situational awareness will tend to atrophy slowly with disuse, and practice will have a crucial role in maintaining it.

6. Situational awareness will be improvable through the acquisition of appropriate skills, by the development of expertise, by a more extensive knowledge base, and by improved accessibility of that knowledge base.

7. Situational awareness will be subject to the formation of habits, may be resistant to new evidence that appears to contradict what is already known, may be biased in the choice of what is relevant to it, and may be influenced, and perhaps overly influenced, by memories which, once recalled, may be treated as more relevant than they are.

8. The learned meanings imposed on what is perceived, and which form an intrinsic part of the perception, will be very resistant to the recognition and correction of any errors that they embody (e.g., when a three figure number has been perceived to be a flight level, it may be very difficult to acknowledge that it is actually a speed or a heading).

9. Situational awareness will be extensively influenced by training, by what is taught, by how it is taught, and by the relevance of what is taught to what is needed.

10. Situational awareness will be influenced by motivation and interest and will furnish opportunities for satisfaction and esteem.

11. Situational awareness will be affected by variations in attention.

12. Situational awareness does not incorporate all stimuli, but only those that can be made meaningful, and items that seem meaningless will not be included in situational awareness, no matter how important they may be.

Fragmentation

If situational awareness is a unitary and unifying concept, it can provide an effective and powerful tool to integrate previously disparate themes, applications and functions, and it can be applied to almost every specialized topic and context within aviation. This universality is its great strength but also its main potential weakness and limitation, for it would be undermined by any fragmentation, and there is a major difficulty in devising and proving means to measure situational awareness without splitting it up. Any attempt to partition or divide it must weaken it as a tool, for situational awareness as a whole is not the sum of its parts.

Automated Aids

The provision of automated aids may not merely change situational awareness, but must change it if the aids are used. The reason is that all aids require new learning of some kind, and situational awareness is a function of learning. All the major proposed forms of computer assistance for air traffic controllers in performing their tasks, and all the intended forms of automation in air traffic control that are envisaged to have some consequences for the controller, must affect situational awareness. The expressed anxieties about some of the consequences for situational awareness of increased air traffic control automation, such as an increased propensity for the controller to lose the picture or reduced controller understanding of the picture, seem to have some justification. The effects on situational awareness are not usually among the given reasons for introducing automation but are among the consequences of it which at best have been foreseen (but often have not been) and are unplanned and unwelcome when they appear.

Incidental Effects of Automation

There are many examples of the incidental effects of the anticipated progressive automation of air traffic control on situational awareness. Among the most significant are the following:

1. Many spoken messages between controllers and between pilots and controllers are being replaced by transponded data which appear automatically not in verbal form but in visual form on the controller's displays as new or updated information.

2. Decision aids are proposed which provide ready-formulated solutions for the controller to accept or reject, with no requirement to work out solutions, to know the reasons for them, or to recognize circumstances which would warrant their revision.

3. Air traffic control is evolving from the hands-on tactical control of each aircraft to hands-off strategic planning of air traffic flows.

4. There will in the future be less qualitative information available to the controller about the trustworthiness of data, and less information actually or potentially available to act as precursors of unsafe acts or to denote actual or incipient failures and the extent of their effects within the system.

5. Although it is probable that in general the greater the human workload is the greater the situational awareness becomes, unless workload approaches levels that may lead to loss of picture, the policy is to reduce workload for other reasons.

6. Paper flight progress strips, which are incorporated into the controller's picture partly through active manipulation and annotation, are being replaced by electronic flight strips where the corresponding functions are fulfilled automatically.

7. Automation renders the air traffic control workspace much less open and observable. The current openness and observability of the air traffic control workspace permits broader and more intensive situational awareness by controllers, colleagues, supervisors, assistants, instructors and others directly concerned with it, allows other controllers to recognize the needs, impending problems, or difficulties of a colleague and to lend effective help, is the basis for judgments of professional competence and respect, allows knowledgeable observers to "read" the traffic, and reveals to each controller how closely the control methods and procedures of colleagues conform with his or her own.

What all the above examples have in common is that every one of the projected changes in air traffic control will tend to reduce rather than increase the situational awareness of the individual controller.

Conclusions

The main conclusion is clear and quite stark. If present plans reach function, future controllers will have much less situational awareness than current controllers. It is necessary to plan now for reduced situational awareness in air traffic control in the future because the safety and efficiency of future air traffic control systems must rely much less on the controller's situational awareness. The only alternative is to preserve current levels of

situational awareness, but this requires positive planning because it will not be achieved serendipitously. There do not seem to be any further practical options.

Acknowledgments

I wish to thank CAHFA for their initiative in holding this conference, and Lloyd Hitchcock for his invitation to take part in it. The paper was written at Embry-Riddle Aeronautical University during a semester there as a Human Factors Consultant, and I am grateful for their support in preparing it. Any credit for the paper is therefore shared with the above, but any blame for the views expressed is all mine.

References

Garland, D.J., Phillips, J.N., Tilden, D.S., & Wise, J.A., (1991). *Theoretical Underpinnings of Situational Awareness: Towards an Objective Measure.* Daytona Beach, Fl:Embry-Riddle Aeronautical University Press. CAAR-15498-91-1.

Garland, D.J., Stein, E.S., Wise J.A., & Blanchard, J.W. (1993). Situational Awareness in Air Traffic Control, A Critical yet Neglected Phenomenon. In Garland, D.J. & Wise, J.A., (Eds.). *Human Factors and Advanced Aviation Technologies*. Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Hopkin, V.D., (1967). Unawareness. *Air Clues, 21, 10,* 342-346. [Reprinted (1975): *Air Clues, 29, 3,* 102-104.]

Reber, A.S., (1985). *The Penguin Dictionary of Psychology.* New York: Viking Penguin.

Taylor, R.M. (Ed.) (1991). *Situational Awareness in Dynamic Systems*. Farnborough, Hants, U.K: Royal Air Force Institute of Aviation Medicine Report No. 708.

Controlling Automation in Future Air Traffic Control: The Impact on Situational Awareness

Daniel J. Garland[1] & V. David Hopkin[2]

[1]Embry-Riddle Aeronautical University
[2]Human Factors Consultant

North American Airlines flight 483 and European Airlines flight 390 are 250 miles apart and on a collision course. Neither flightcrew is aware of this. The air traffic controller responsible for North American 483 is not aware of European 390 because the flight is not present on the radar-display. However, the situation is under the control of an "intelligent" computer aid which discovers the conflict, quickly calculates the best solution, and via computer data links, automatically transmits the data to North American 483. Subsequently, North American 483's flight management system receives and initiates the "intelligent" aid's solution: "Descend to 28,000 feet. Turn left heading 248." Later when the two aircraft pass, they are safely separated by several miles.

Introduction

Today, the above scenario is moving further from science fiction and closer to reality. International aviation research, development, and implementation efforts are evolving towards a more automated international air traffic system. The progressive introduction of automated assistance in the air traffic system is warranted by dramatic impending increases in the amount and complexity of air traffic. However, the impact of such automation on the situational awareness of the individual controller is uncertain. Automated aids have great potential to enhance the performance of the controller and the safety and efficiency of the entire air traffic system; yet if these aids also render the control environment sterile, their presence would undermine the necessary and sufficient cognitive processes needed to acquire and maintain controller situational awareness. As the deliberately ambiguous title of this article suggests: Will we control the automation in future air traffic control systems, or will the automation control us?

During the last 20 years, U.S. air traffic has increased from 390,000 to about 900,000 scheduled flights per month. In 1991, commercial airlines reported nearly 300,000 delays of 15 minutes or longer and in 1990, at the ten busiest airports, airlines experienced 590,000

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

hours in delays. Second only to weather, overcrowded airspace and airports are reported as a leading cause of delays (Langreth, 1993). The Federal Aviation Administration estimates that in the United States, delays related to air traffic problems result in economic losses of over five billion dollars per year. These losses are expected to exceed ten billion dollars per year by the year 2000, if no changes are made (Wise, Hopkin, and Smith, 1991).

The need for improved air traffic systems is widely recognized throughout the world. As a result, practically every industrialized country is trying hard to improve the performance of its air traffic system. Most are developing and/or installing advanced automated aids to be used in both operations and maintenance. The international aviation community is actively promoting advanced computer systems and software that are intended to enhance air traffic system safety and performance. The following are examples of a few international aviation research and development efforts in air traffic control (ATC).

- *Automated air traffic conflict resolution technologies.* These technologies (e.g., AERA, USA; CORES, Canada; ASTA, Eurocontrol) are capable of detecting and resolving en-route conflicts minutes before they occur. From the flight information about the aircraft in a given area, flight course projections are calculated as much as 15 to 25 minutes into the future. When a conflict is predicted, solutions are determined and offered to the controller, who then decides what to do. In their most advanced form, conflict resolution technologies would determine the solutions and communicate them to the flightdeck via computer data links, without informing the controller.

- *Automated aids for prediction of optimal timing and sequence of aircraft into an efficient arrival stream.* Such technologies (e.g., CTAS, USA; COMPAS, Germany; MAESTRO, France) check radar returns for aircraft as much as 200 miles away. Recommendations are then presented on when and in what order aircraft should land. In addition, as aircraft approach the airport, recommendations are made on the spacing between aircraft, on speed changes, and on when each aircraft should initiate turns on descent. Essentially, such technologies advise controllers in maneuvering aircraft on approach for landing.

- *Air traffic control tower voice-recognition computer systems.* These would process and comprehend controllers' voice commands to the flight crew and compute where each aircraft ought to be according to those commands.

- *Computer data links between ground and flightdeck computers.* These are capable of transmitting and receiving everything from routine preflight clearances to up-to-the-minute weather information, but especially information on the position and state of each aircraft. In the future, all ground-to-air communications may be conducted via computer data links, making voice communications obsolete.

- *Electronic flight data displays (i.e., electronic flight strips).* These will replace current paper flight strips, providing up-to-minute controller-entered or system generated electronic flight data about each aircraft in a given region.

In addition, advanced automated aids are being developed to assist in processes such as (a) the training of air traffic controllers (e.g., IATCTS, USA), (b) decision making aids based on

a problem-driven information filtering for ATC (e.g., SMARTFLOW, USA; ERATO, France), and (c) air traffic management (e.g., CINTIA, Belgium).

During the next decade and beyond, these advanced automated aids and other new technologies will revolutionize the international air traffic system and the controller's job. The implementation of new automated aids, while beneficial, will not resolve all the problems of air traffic control. In fact, new problems are inevitable (Garland and Wise, 1993; Wise and Debons, 1987; Wise et al., 1991; Wise, Hopkin, and Stager, 1993). For example, while the implementation of automated aids (e.g., expert systems, memory aids, decision aids) is intended to enhance a controller's situational awareness, these aids may instead impede or prevent the necessary and sufficient cognitive processing required for effective situational awareness of the air traffic environment (Garland and Stein, 1992; Garland, Stein, Blanchard, and Wise, 1992; Garland and Wise, in press; Hopkin, 1992; Stein and Garland, 1993).

Dramatic system changes in air traffic control automation will not only replace existing ATC technology and equipment, but will also fundamentally change the way in which air traffic controllers conduct their job. Air traffic control is gradually evolving from "hands-on" tactical control of each aircraft to "hands-off" strategic planning and management of air traffic flows. Consequently, there is a concern in the ATC community, that progressive automation may impose requirements on the controller that are incompatible with the way the controller processes information.

The cognitive requirements of air traffic control involve the processing of a great volume of dynamically changing information. Cognitive processing of flight data (i.e., call sign, aircraft type, sector number, planned route, assigned speed, heading, altitude, time over posted fix, etc.) is crucial to virtually every aspect of a controller's performance. It is essential for the controller to manage information resources in such a way that accurate information is available when needed. The ease with which information (e.g., flight data) is processed and remembered depends on how it is displayed and how the controller interacts with the information. The dramatic changes to information display and analysis resulting from ATC automation may influence the processing of information, potentially affecting ATC performance and situational awareness.

Situational Awareness

A Unifying Concept

It is not yet ten years since situational awareness made its appearance as a widely used concept, but in the interim it has become a popular concept which clearly meets a need. Speculations on the reasons for this may help to explain its usefulness. Perhaps situational awareness is a recent concept because it used to be given a variety of other names in various contexts. If this is so, the potential value of situational awareness as a concept begins to emerge. It is a unifying concept, and perhaps a universal one, for it provides a means to reveal commonalities across contexts and applications, and also to permit practices and constructs to be brought together and compared that previously had been used and even named differently but that now seem to have much in common. Situational awareness is also

a unifying concept in another important respect: It treats as an indivisible whole a notion previously partitioned into various aspects that were addressed separately.

If situational awareness is a unitary and unifying concept, it can provide an effective and powerful tool to integrate previously disparate themes, applications and functions, and it can be applied to almost every specialized topic and context within aviation. This universality is its greatest strength, but also its main potential weakness and limitation, for it would be undermined by any fragmentation. There is a major difficulty in devising and validating means to measure situational awareness without splitting it up. Any measures which do divide situational awareness must weaken it as a tool, for situational awareness as a whole is greater than the sum of its parts.

"Situation" or "Awareness"

Many people have had the disconcerting experience, particularly while driving a car along a featureless but familiar road with little traffic, of suddenly becoming aware that they have no recollection of the last few miles they have driven. Does this experience constitute a loss of situational awareness or does it not? The question is neither trivial nor trite. It points to an anomaly in the literature on situational awareness and in the nature and measurement of the phenomenon itself. Is the emphasis on the situation, or on awareness of the situation?

If situational awareness primarily concerns the situation, then it is measurable in terms of performance, and the fact that any passenger in the car usually reports no noticeable change in driving performance during this gross lapse of attention suggests that situational awareness remains unimpaired because performance of the driving task remains intact. If situational awareness primarily concerns awareness, then in this driving condition situational awareness has been lost because awareness has been lost, but subjective measures are necessary to establish this. Different measures, in this instance performance and subjective measures, can yield opposing evidence on whether situational awareness exists. In the literature, some of the claimed measures of situational awareness are of performance and others are subjective. Trying to have this both ways engenders confusion about situational awareness itself, and results in potentially anomalous or contradictory findings. What sort of measures could prove that two people had the same situational awareness?

To further pursue the example of driving a car, neither performance nor subjective measures would suffice to examine situational awareness. Performance measures are probably so insensitive that they cannot even distinguish reliably between the presence or absence of this gross lapse in awareness. Subjective measures are not much better: It is necessary to rely on them for evidence of the lapse of awareness, but they are useless thereafter in dealing with the phenomenon if its essence is that nothing occurring during it can be recalled. This relates to automaticity or automatization, featured in many theories or descriptions of situational awareness, sometimes claimed as essential for situational awareness and sometimes as a kind of negation of situational awareness.

In considering situational awareness in air traffic control, perhaps the first issue to be addressed is whether the human factors work has been independent or, having far fewer resources, has tagged along after the work on cockpits and borrowed, adapted or appropriated studies done on flight decks. This does not seem to have happened; rather the work in air traffic control has paralleled cockpit studies independently using different terminology, notably the controller's "picture."

Lapses of attention of the kind that occur during car driving are not commonly reported in air traffic control. This could be either because they are genuinely rare or because common prudence suggests that a controller might be unwise to publicize them. Air traffic control does however pose a further question, the answer to which seems to depend crucially on this apparent dichotomy between the situation and awareness of it. The question is: Can the air traffic controller have too much situational awareness?

Too Much Situational Awareness?

While the preponderant view seems to be that it is not possible to have too much awareness, it is certainly possible to have too much situation that the controller must be aware of. This is the essence of the problem when the controller "loses the picture." This occurs when the amount of information exceeds the controller's capability to maintain it as an integrated, coherent, and meaningful entity for the purposes of the control of air traffic. It is associated with heavy traffic, at a time when there is generally more information to be integrated. Its onset can be sudden, but is often preceded by a period during which the controller becomes increasingly worried about the possibility of sudden loss of picture. If it does occur, the operational consequences can be very serious. Once the picture has been lost, the controller can seldom recall it in its entirety but has to rebuild it painstakingly aircraft by aircraft. This process often requires the systematic re-committal to memory of the details of each aircraft under control, one at a time, with the realization that if this reconstruction process is successful, not only will the picture have been restored but so will the same conditions which previously led to loss of picture.

Whether there can be too much situational awareness has profound consequences for measurement. If there cannot be too much situational awareness and it has no maximum, then no valid measures can be devised that purport to deal with proportions or percentages of it. But if it has a maximum or optimum, such types of measures become feasible, at least in principle. In the former case it is never possible to prove that full situational awareness has been achieved. In the latter case, this may be possible.

Cognitive Skill Acquisition and Situational Awareness

The cognitive requirements for the development, maintenance, and enhancement of situational awareness involve the processing of a great volume of dynamically changing and interacting information. Cognitive processing of flight and airspace information is related to every aspect of a controller's performance. It is essential for the controller to manage available information resources in such a way that accurate information is present or obtainable when needed. The ease with which information is processed and remembered depends on how it is displayed and how the controller interacts with the information. Controllers have to assess situations immediately and act at once according to their current knowledge of each situation.

The cognitive processes used by air traffic controllers to interact with ATC equipment and to perform the required tactical operations are fundamental to superior situational awareness.

These processes involve the ability to extract, integrate, assess, and act upon task-relevant information. Those who are able to acquire and maintain a high level of situational awareness, essentially have a distinct cognitive advantage.

Air traffic control situations, which are punctuated by varied demands and information loads, require continual information acquisition and assessment in order to achieve and maintain adequate situational awareness. Such time-intensive, labor-intensive situations often necessitate the use of highly practiced automatic response patterns and the employment of the full range of cognitive structures to process associated data and knowledge that provides context and meaning to the air traffic situation.

Superior situational awareness is a product of cognitive skill refinement. The integrity of situational awareness is founded in domain-specific cognitive skills, and much of situational awareness consists of knowing how task operations work and when to use them. The acquisition of this knowledge is an important step in understanding how superior situational awareness evolves and is maintained.

In considering skill-acquisition processes, Fitts (1964) has suggested that skills are developed in stages, starting with a cognitive stage, followed by an associative stage, and then a final autonomous stage. Anderson (1982) has modified Fitts' stages to describe cognitive skill acquisition. The initial stage is declarative knowledge acquisition, followed by knowledge compilation in which declarative knowledge is translated into procedures. Then there is a final procedural stage during which procedures are characterized as autonomous.

In addition to Fitts' (1964) and Anderson's (1982) theories, several other theories have been proposed to account for the processes used in skill acquisition (e.g., Schneider and Shiffrin, 1977). Although a review of these theories and others is beyond the purpose of this article, the primary processes of each position are qualitatively similar.

The initial knowledge acquisition phase is generally characterized by slow, deliberate processing, with strong demands placed on the cognitive system (i.e., working memory), and a great deal of attention given to the formulation of production strategies and the understanding of the task. With extensive and consistent practice, cognitive processing is reduced and strategies are fully formulated, resulting in increased speed and accuracy of performance. In the skill-acquisition phase, production strategies are further refined with use. The final skill-refinement phase is characterized by effortless automatic processing and exceptional cognitive skills (e.g., automaticity, skilled memory, dynamic working memory), and in many task environments this phase remains quite elusive. This fine tuning and automatizing of task-specific cognitive skills is seen as a prerequisite for superior situational awareness. It is recognized that different levels of situational awareness may be achieved dependent upon the level of cognitive skill acquisition attained. However, without the development of increasingly specialized production rules to apply to any of the possible problem types which may be encountered, superior situational awareness will be unrealized.

Situational awareness therefore is learned, and there can be gross individual differences in actual situational awareness and perhaps also in the potential for situational awareness in different individuals. Several consequences follow from this, which can be stated as hypotheses:

1. Different controllers, given the identical displayed information, can have different situational awareness, so that the situational awareness of one controller might not include everything that another controller would assume it contained.

2. Because situational awareness depends on learning, it should to some extent be explicable and predictable in terms of learning concepts and theories, provided they are valid.

3. Situational awareness, as the product of a learning process, can be built, extended, developed, entrenched and reinforced.

4. Situational awareness, also as the product of a learning process, can be fallible, incomplete, distorted, subject to error, or forgotten.

5. The capability to acquire situational awareness will tend to atrophy slowly with disuse, and practice will have a crucial role in maintaining that capability.

6. Situational awareness will be improvable through the acquisition of appropriate skills, by the development of expertise, by a more extensive knowledge base, and by improved accessibility of that knowledge base.

7. Situational awareness will be subject to the formation of habits, may be resistant to new evidence that appears to contradict what is already known, may be biased in the choice of evidence that is relevant to it, and may be over-influenced by particular memories which may be treated as more relevant than they are.

8. The learned meanings which form an intrinsic part of the perceptual structure will be very resistant to the recognition and correction of any errors that they embody (e.g., when a three-digit number has been recognized as a flight level, it may be difficult to acknowledge that it is actually a speed or a heading).

9. Situational awareness will be extensively influenced by training, by what is taught, by how it is taught, and by the relevance of what is taught to what is needed.

10. Situational awareness will be influenced by motivation and interests, and can furnish opportunities for job satisfaction, self-esteem, and the esteem of others.

11. Situational awareness can be influenced by attentional factors.

12. Situational awareness does not incorporate all stimuli, but only those that can be made meaningful, and items that seem meaningless will not be included in situational awareness, no matter how important they may be.

The Controller's "Picture"

In air traffic control, situational awareness seems to correspond quite well with the concept of the controller's picture. This picture is sometimes construed as an example of a mental model. Although the picture includes the controller's mental model, it is not confined to it,

being a more dynamic entity than most mental models in that it incorporates changing states and their consequences. In some respects, both the concept of the picture and its nature seem specific to air traffic control. Typically, the controller of a sector (a geographical region of airspace containing en route traffic at high flight levels) acquires the "full picture" over a period of about fifteen to twenty minutes because by then the controller knows the full history of all the traffic while it has been under his or her direct control. However, an adequate picture for controlling the traffic can be built much more rapidly, and this occurs whenever one controller hands over the control responsibility for a sector to another controller at the end of the work shift, or whenever a supervisor "reads" a controller's traffic and adds an annotation to a controller's paper flight progress strip for a particular aircraft.

A controller's picture may be specific to a given situation (e.g., VFR traffic) or more global to the entire task domain (e.g., the entire flight sector). It may or may not include abstractions concerning functional relationships, operating guidelines, and systems goals and objectives (Mogford, 1991; Norman, 1986; Rasmussen, 1979; Wickens, 1992; Wilson and Rutherford, 1990). Research on mental models and conceptual structures in the air traffic control environment is disappointingly limited (see Mogford, 1991, for a review). However, the research that is available does suggest a connection between a controller's picture and understanding of, and memory for, the traffic situation (e.g., Bisseret, 1970, 1971; Means, Mumaw, Roth, Schlager, McWilliams, Gagne, Rosenthal, and Heon, 1988; Moray, 1980; Whitfield, 1979). General conclusions of these studies are that skilled controllers, in comparison to less skilled controllers, use their picture as a "supplementary display" in order to enhance memory for aircraft, and that the quality and functionality of the controller's picture are directly related to ATC expertise.

To anyone with no knowledge at all of air traffic control, the controller's workspace is meaningless and mystifying. This is because it contains no information about what it is, what it is for, or how and why it could or should be used. Its meaning depends therefore on what is known as well as on what is portrayed. Only the controller can make sense of the portrayed information in terms of a picture of the air traffic. Therefore, only a controller could have full situational awareness in an air traffic control workspace. The naive occupant of the workspace would see a room, furniture, meaningless information displays, and keys with unknown functions. Does this constitute a superficial form of situational awareness, or is situational awareness of an air traffic control workspace restricted in principle to those with professional knowledge of air traffic control? If the latter is required, according to what criteria could the trainee controller be said to possess sufficient knowledge to have situational awareness, and what kind of knowledge would that be? The naive observer has a different kind of situational awareness or no situational awareness at all, and can have no insight into the level of complexity of awareness that the situation is capable of. At present, air traffic control is not a self-teaching environment, although whether it should become more self-evident and transparent to its users in the future is currently a debated issue.

A better understanding of the controller's picture is needed as ATC systems become more automated, forcing the controller into ever increasing levels of supervisory control. There also needs to be a better understanding of how increased computerization of ATC tasks influences the development of the controllers' picture and its potential supporting influence on controller situational awareness. An understanding of the controller's picture may suggest appropriate forms of automation for controller training and memory aids since such aids must interact with the cognitive processes of the controller to be effective (e.g., Hollnagel and Woods, 1983; Moray, 1988). The organizational format of the data must be compatible with the operator's conceptualization of the data. Data in an inappropriate format may be

impossible to incorporate into the controller's picture, may be distorted by re-formatting processes, and/or may be flawed by errors made in re-coding the data to render it more compatible with the controller's picture.

The Impact of Automation on the Future ATC System

Early forms of automation in air traffic control were applied to the gathering, storage, compilation, condensation, retrieval, and presentation of data for use by the controller. These automated forms have evolved to the point where voice communications between controllers and pilots are being progressively replaced by automatically transponded data, an increasing trend in the future with the advent of computer data links. These data contain quantitative information, but have no qualitative data comparable to pace, hesitancy, pauses, sequencing, rigidity of message formats, tone of voice, accent and the like, which are used by controllers to make judgments about the competence, confidence and experience of pilots and used by pilots to make comparable judgments about controllers. The effects of removing qualitative information on situational awareness and resultant safety are unknown. They need to be ascertained beforehand since if some qualitative information proves to be vital for situational awareness and safety, a substitute for it will have to be found.

Most forms of ATC automation now planned are more advanced than the original "data-crunching" forms, and are applied to assist more cognitive and traditionally human functions such as scheduling, prediction, problem solving, and decision making. Data crunching aids are generally acceptable to controllers and are found to be helpful. Controllers are much more wary of aids which impact directly on their skills and their responsibilities, especially if modifications to their skills and responsibilities appear to be needed in order to make use of the automation. These forms of automation can be very helpful if correctly used, and they would not be introduced unless they were known to be reliable and safe. Generally, the decisions of the automated aids are indeed the best decisions, and the automated solutions to problems are safe and may be near optimum. However, this carries implications of complacency which have been recognized, and of over-protectiveness which have not.

When an automated aid is always right and the controller has to choose or reject the offered solution, the tendency is gradually to learn to trust it. Corresponding human information processing as a backup to check that all is well generally becomes a vestige of its former thoroughness since it is no longer needed, and may disappear altogether. If the controller's task includes the acceptance by pressing a key or similar simple action of what the computer has formulated, others such as supervisors, colleagues, or managers have no immediate means to judge how competent the controller is or whether the controller understands what is happening. A controller's skill could dramatically degrade and the system would protect him or her and disguise human incompetence and inadequacy.

A system may present a series of solutions to a particular problem in the order of preference of the computer for the controller to choose the preferred one. This can be successful provided that in all circumstances the computer will present at least one solution. A misapplication of automation occurs if there are circumstances when there is no automated solution to a problem within the rules. This means that the automation could be inherently dangerous and also implies that any solution formulated by the controller must override some

rules and may therefore be declared invalid by the computer, since otherwise the computer would have presented the solution. Thus the controller is put in a position of having to override and violate the rules, and there may not even be adequate means of doing so. Such problems must be avoided. The computer must advance solutions while they are still available, and not delay until there are none. Certain forms of aid being considered now have this problem inherently within them.

Future air traffic control automation appears to offer the distant prospect of replacement of the human controller entirely. It becomes possible, for example, to detect a potential conflict between two aircraft, formulate a solution, evaluate that solution, implement the solution, and check that the conflict has been resolved; and to do all this automatically without any reference at all to the controller or pilot. This raises some difficult issues, one of which is where the legal responsibility for any failure lies. Coupled with this is the likely insistence that if the human must carry the legal responsibility, then the human controller must have the means and the knowledge to intervene in order to exercise that responsibility. This intervention negates in principle such a fully automated system. Therefore, highly automated systems lead to a new generation of problems in the partitioning or reconciliation of human and machine functions, particularly in relation to the responsibilities. The traditional allocation of functions to human or machine becomes an invalid approach to the problem. An envisaged form of computer assistance might be to inform the machine about what the controller is trying to achieve so that the controller can enlist the support of the machine in achieving the human objectives. Currently this can seldom be done.

The technologies that are adopted in air traffic control are not adopted for human factors reasons. Whether they are helpful in human factors terms therefore is somewhat arbitrary. The earlier technologies of better data gathering, of radar, of information processing and presentation, and of automated conflict detection or aids to show that an aircraft was departing from its prime route, for example, were universally helpful and this was acknowledged. More recent technologies are less immediately compatible with human factors requirements because they impinge so closely upon human skills and responsibilities and still require the human to adapt to them rather than constitute forms of assistance for the human. This tends to be true of problem-solving, decision-making, and prediction aids for example. The machine is built to be very helpful in detecting a problem but it may be difficult for the controller to determine which options have been considered, why certain options have been rejected and how far the system has been planning ahead in proposing a particular solution. For example, initial forms of automated conflict resolution may resolve each conflict between two aircraft without discriminating between some solutions which would precipitate a further conflict later and other solutions which would not. The problem of trust arises as more complex and cognitive forms of automated assistance are provided, and it is a two-way problem. The controller must learn to judge the degree of trust that is appropriate for the forms of computer assistance provided, but the machine must be programmed to accept some human actions but to veto, challenge, or require confirmation of others.

More advanced forms of technology, including data links and satellite-derived information, raise issues of what the human roles ought to be in relation to these data. They tend to be quantitative rather than qualitative and they have to be trusted, with no means of checking whether they should be or not. An implication of this is that it may be very difficult for the user to tell if they have been degraded or if they have failed, since it is not evident what the cues would be or whether there would be any cues at all from the limited summary of information available to the controller.

The justification for automation has always been associated with problems of workload. From the outset, one of the stated reasons for automated assistance has been to reduce the amount of work of the controller. The only way in which air traffic control systems can handle larger numbers of aircraft within the same airspace without increasing the number of controllers, which is counterproductive as a solution because of the extra coordination and liaison entailed, is to use automation so that each controller has to spend less time in handling each aircraft. The implication is that it is self evidently advantageous, and indeed essential, to reduce controller workload. However, this is not necessarily the case, particularly if the removal of certain routine functions means that the controller has to process less information or the same information at less depth. As a result, the controller becomes aware of a loss of understanding and a reduced appreciation of the total air traffic picture.

Maintaining Situational Awareness in the Future ATC System

Some simple forms of computer assistance were introduced into air traffic control long ago. They exemplified that forms of computer assistance can differ in their implications and acceptability. For flight progress strips, an early form of computer assistance was the automated printing of them from the flight plan information that had been filed. This saved clerical work and rendered the delivery of flight strips more reliable and predictable to the controller if this was properly arranged, and made little difference to actual controller tasks except to ensure that the information on the strips was uniformly readable. This form of assistance was therefore accepted and welcomed. On the other hand, the initial introduction of radar displays was a radical departure from the forms of information that had been previously available, though the technical development was crude by modern standards and most of the matching between human and machines had to rely on human adaptability because the machine was inflexible. Eventually, further forms of computer assistance made the radar displays much more compatible with human needs because identity, altitude, route and other kinds of information were presented in the form of labels. The introduction of radar displays initially met with considerable resistance and wariness on the part of controllers. It was not so much that the information was not useful because it was in fact very useful for air traffic control, but that it required different cognitive procedures to assimilate it, to understand it, and to use it. Controllers who had no previous experience of controlling on flight strips only accepted radar displays willingly and found them easy to use and an essential and welcome form of computer assistance. All the controllers accustomed to thinking in terms of flight strip categorization found that radar information was not immediately compatible with this kind of thinking but required radical changes in their thought processes and in the ways of assimilating and integrating data in order to use it. The kind of picture of air traffic built up from flight strips is not the same as the kind of picture built up from radar displays, and some of the earliest controllers never really did learn entirely to stop thinking in terms of flight strip information and its categories.

This early example therefore demonstrates one of the problems at the heart of the introduction of forms of computer assistance into air traffic control. If the proposed form of assistance is compatible with existing thought processes, then there need be no major problems in obtaining the expected benefits from it, in integrating it with existing forms of

information, and in making it acceptable, provided that the actual form of computer assistance is made as compatible as possible with human requirements. But if the introduced forms of computer assistance require different kinds of cognitive processing, different ways of thinking and the discarding of traditional methods and skills, then this introduces problems of both efficiency and acceptability. This does not mean that they are not beneficial or that they could not be successful, but it does mean that these benefits will not accrue without recognition of the need for a re-matching of the new forms of computer assistance with existing cognitive processes. In general, the more cognitive the forms of assistance are and the more concerned with higher mental processes which are mainly cognitive, the more this problem of cognitive compatibility of the new with existing thought processes arises, and the more essential it is to identify the nature of the problem and the range of possible solutions beforehand so that the expected benefits of the computer assistance actually materialize.

The above concerns regarding the desirable common functionality between, the present system and the future system are perhaps most dramatic with the near-term implementation of electronic flight strips. Paper flight progress strips, with their functions of active manipulation, annotation, and incorporation into the controller's picture, are being replaced by electronic flight strips, where the corresponding functions are fulfilled automatically. The full impact of such a radical change on the controller's processing of flight data is unknown. There is the real possibility that such change will fundamentally influence the cognitive performance and resultant situational awareness of controllers.

Paper flight strips have become a fundamental part of air traffic control. Hopkin (1991a) states that the paper flight strip acts as an information display, notepad, memory aid, history, and record of actions. Hopkin, in summarizing work on flight strip functionality (e.g., Harper, Hughes, and Shapiro, 1989; Jackson, 1989), indicates that the actual use of the paper flight strip has exceeded its original purpose as a memory aid. He notes that the quest for an electronic replacement for paper flight strips has revealed that paper flight strips are a more complex and powerful tool than was originally believed, with more flexible functionality. For example, the process of physically manipulating the paper flight strips (e.g., inserting or removing them from the strip bay) facilitates the controller's awareness of the physical interrelationships between aircraft (i.e., the creation and revising of the traffic picture). Additional actions, such as sorting the paper flight strips or "cocking" them on the strip board, further facilitate the controller's understanding and memory for the information displayed. Controllers have developed unique ways of sorting and marking that seem to work effectively for each individual.

Further, paper flight strip markings or notations give them their notepad character. Jackson (1989) notes that "it is possible that when controllers read information from a strip in their own handwriting they do not only interpret and comprehend the content, they also remember the previous act of writing it and, perhaps more importantly, the reasons why a particular course of action had been undertaken (e.g., 'Now why did I do that?')" (p. 5). This observation is supported by psychological research on retrieval cues (e.g., Tulving and Thomson, 1973) and research that demonstrates that subject-generated memory aids (i.e., hand-written notes, annotation of to-be-recalled items with a self-generated icon) facilitate memory retrieval (Intons-Peterson and Fournier, 1986; Lansdale, Simpson, and Stroud, 1990).

Research on memory for action events has focused on memory for past activities (e.g., Koriat, Ben-Zur, and Sheffer, 1988). A consistent and general finding of these studies is that memory for performing a task is superior to memory for verbal materials, due to the beneficial effects of motoric enactment. That is, the process of physically performing a task

seems to enhance the encoding of and subsequent memory for the task. The superior memory for performing tasks "has been generally attributed to their multimodal, rich properties, assumed to result in richer memorial representations than those formed for the verbal instructions alone" (Koriat, Ben-Zur, and Nussbaum, 1990).

These results are particularly relevant when discussing the impact of progressive automation on ATC systems and the potential human factors consequences. Several researchers (e.g., Hopkin, 1988a, 1988b, 1989, 1991a, 1991b; Narborough-Hall, 1987; see Wise and Debons, 1987; see Wise et al., 1991) have suggested that routine task performance facilitates controller tactical operations (e.g., the understanding of and the memory for traffic situations). Hopkin (1991a) asserts that physical interaction with the flight progress strip is fundamental to a controller's memory for immediate and future traffic situations. The impact of automating routine controller tasks (e.g., physically marking the flight strips) that facilitate the development of the controller's picture is unclear.

Several researchers have identified the significant cognitive value of paper flight strips in preparing for future actions (Hopkin, 1989, 1991a; Vortac, 1992a, 1992b). Paper flight strips can represent the history of actions, goals, intentions, and plans of pilots and controllers if controllers are using them effectively. These functions are elaborated in the following controller interview extract (Harper, Hughes, and Shapiro, 1989):

> It's a question of how you read those strips. ... An aircraft has called and wants to descend, now what the hell has he got in his way? And you've got ping, ping, ping, those three, where are those three, there they are on the radar. Rather than looking at the radar, one of the aircraft on there has called, now what has he got in his way? Well, there's aircraft going all over the place, now some of them may not be anything to do with you, ... your strips will show you whether the aircraft are above or below them, ... or what aircraft are below you if you want to descend an aircraft, and which will become a conflict. You go to those strips and you pick out the ones that are going to be in conflict if you descend an aircraft, and you look for those on the radar and you put them on headings of whatever, you find out whether those, what those two are--which conflict with your third one. It might be all sorts of conflicts all over the place on the radar, but only two of them are going to be a problem, and they should show up on my strips (p. 9).

This interview extract provides a good example of the role flight strips play in assisting information processing and its significance in planning future actions. Harper, Hughes, and Shapiro (1989) point out that paradoxically, the "moving" radar screen is from an interpretative point of view relatively static, while the "fixed", "hard copy" strip is interpretatively relatively dynamic. For ATC tactical operations, planned actions are the purview of flight progress strips, and past actions are reflected in feedback on the radar and flight strip markings (Vortac, 1992a, 1992b).

Directly related to memory codes, particularly motoric encoding, is a robust memory phenomenon known as the "generation effect" (Slamecka and Graf, 1978). Simply stated, the generation effect refers to the fact that information actively and effortfully generated (or information which you are actively involved with) is more memorable than passively perceived information. The essence of this memory phenomenon is expressed in the sentiment that there is an especial advantage to learning by doing, or that some kind of active or effortful involvement of the person in the learning process is more beneficial than merely passive reception of the same information (Slamecka and Graf, 1978).

The generation effect has direct relevance to ATC tactical operations, where the active integration of the controller's information processing capabilities with the relevant support systems (flight progress strips, radar, etc.) is fundamental to the integrity of the understanding and memory of the traffic situation. Means, Mumaw, Roth, Schlager, McWilliams, Gagne, Ronsenthal, and Heon (1988), using a "Blank Flight Strip Recall Task," demonstrated that controllers' memory for flight data is a function of the level of control exercised. Their data indicated that memory for flight information of "hot" aircraft, which required extensive control instructions, was significantly better than memory for flight information for "cold" aircraft, which required little controller intervention (e.g., overflight).

The foregoing discussion suggests the importance of a direct manipulation environment (Hutchins, 1986) for ATC. Such an environment seems essential to maintain and potentially enhance the integrity of ATC situational awareness. In an analysis of flight progress strips, Hopkin (1991a) indicates the cognitive significance of flight strip manipulation.

> Strips help the controller to organize work and resolve problems, to plan future work, and to adjust current work in accordance with future plans. The physical act of transferring the strip from the pending to the active bay or assuming control responsibility for an aircraft involves a recapitulation and review of knowledge and previous decisions. This process reinforces the picture of the traffic as a whole, and the details recalled about each aircraft. The physical action in moving a strip aids memory of its contents, of its location on the board, and of why it is there. Writing on flight strips seems more memorable than watching the automatic updating of information"...(on electronic flight strips) (p. 2).

Hopkin (1991a) further comments:

> ...whatever form electronic flight strips take, it is essential to define beforehand all the functions of paper flight strips, in order to discard any unneeded functions deliberately and not inadvertently, to confirm that familiar essential functions can still be fulfilled electronically, and to appreciate the functional and cognitive complexity of paper flight strips. Electronic flight strips have major advantages in compatibility with computer-based air traffic control systems, but their compatibility with human roles is less obvious, requires positive planning, and depends on matching functions correctly with human capabilities (p. 3).

Manipulative control actions, both routine and strategic, required by the controller appear to be fundamental for tactical operations and situational awareness. An obvious concern for current and future ATC systems is optimizing controllers' direct manipulation of the system. This optimal manipulation seems fundamental for ATC system performance.

The assumption underlying the implementation of new technologies (e.g., electronic flight strips) is that with the automation of functions which were once allocated to human control, the processing resources of the controller will be freed to deal more effectively with other required aspects of the system. This is an assumption which is not necessarily proven or possible. It implies that the other aspects can be dealt with more effectively; that is, they are susceptible to improvement if the controller can devote more time to them. It also implies that it is practical to redirect human resources to whatever alternative functions are designated, which may also be an unwarranted assumption. However, while the use of new technologies may be essential in order to deal with the ever increasing information processing

demands of the ATC system, the long-term performance implications of extended use of the new technologies on human performance are largely unknown.

It is not a question of whether situational awareness will change as a result of the introduction of automated aids into the air traffic system, but a question of how much will it change. This is because all aids require new learning of some kind, and situational awareness is a function of learning. All the major proposed forms of computer assistance for air traffic controllers in performing their tasks, and all the intended forms of automation in air traffic control that are envisaged to have some consequences for the controller, must affect situational awareness. The expressed anxieties about some of the consequences for situational awareness of increased air traffic control automation, such as some increased propensity for the controller to lose the picture or reduced controller understanding of the picture, seem to have some justification. The effects on situational awareness are not usually among the given reasons for introducing automation, but are among the consequences of it.

Where do we go from here?

As advanced control facilities come on-line with controllers handling requests never before honored, the question of system capacity is sure to arise. How will the implementation of new automated aids influence system capacity? Controller situational awareness should play a role in the answer. System designers should consider the advanced automated control facility to be a concept inclusive of active human involvement. Controllers have and will undoubtedly continue to have input into the design of control facilities, but are they providing this expertise based on today's system of highly active, hands-on controlling or with the realization that all the data on the scope will be managed by a computer? In order to improve situational awareness, the advanced control system design process must embrace the concept that the human controller will continue to be the decision-maker, assisted by a computer with two primary roles: (1) flexible information supply, and (2) monitoring of overall systems operations. Automated systems can provide a high degree of information in a relatively simple format if designed with the user's flexibility in mind. Fundamental to improving potential situational awareness is the flexibility of the facility to assist the controller in the decision-making process, thereby maintaining the controller's picture (Garland, Stein, Blanchard, and Wise, 1992).

All the projected changes in air traffic control will tend to reduce rather than increase the situational awareness of the individual controller. This implies that the achieved or achievable levels of situational awareness of the controller are an incidental consequence of forms of computer assistance and other system changes introduced for reasons unconnected with situational awareness. This is clearly an unsatisfactory way to proceed, and a positive policy on the desirability or otherwise of situational awareness in future air traffic control systems is needed.

As long as controllers have a legal responsibility for the safety of air traffic under their control, the means to exercise that responsibility have to be provided. Among them is a requirement to maintain continuous awareness of the current and pending air traffic control situations, which implies situational awareness at least to the level required for the legal responsibilities. This awareness cannot be sustained passively but entails active involvement

in control loops and functions. This human involvement is thus an essential condition, in its own right, for the human to exercise responsibilities and respond to emergencies. It should therefore be achieved without prejudice to the forms of computer assistance in use, which are provided for other reasons. If situational awareness can only be achieved by some overlapping or duplication of human and machine functions, this should still be done, since the aim is to optimize the situational awareness in relation to the defined requirements for it. The opportunity should be taken, if duplication of functions is entailed, to discover the potential benefits in the form of safety checks through human and machine comparisons or crosschecks. However, the implicit assumption in the traditional allocation of functions that a function fulfilled by the machine should not also be done by the human, and vice versa should be challenged. The essential point is to identify the requisite level of situational awareness in advance, and aim to achieve it optimally in its own right, and not prevent achievement of the optimum by arbitrary assumptions about what the human-machine relationship should be.

Conclusions

For several decades, there has existed an implicit philosophy of automation which has adopted the assumption that automation is always appropriate. This philosophy has been based, in part, on the availability of increasingly sophisticated and advanced technological innovations, the need to reduce human workload, the need for increased system safety and efficiency, and perhaps, primarily on the assumption that human operators are at their best when doing the least amount of work. While automated systems have provided substantial benefits to the aviation community, the human factors consequences of flawed automation practices are well known.

Any philosophy of automation should be based on an understanding of the relative capabilities of the controller in the system, and the circumstances under which automation should (and should not) assist and augment the capabilities of the controller. What is needed is an approach which has a better philosophical base for what automation seeks to achieve and a more human-centered approach in order to avoid the most adverse human factors consequences of automated systems and provide a better-planned progressive introduction of automated aids in step with user needs. Such a comprehensive, scientifically-based design philosophy for human-centered automation must be developed in order to avoid inevitable "one step forward and two steps backward progression" (Garland, 1991).

The main conclusion is clear and quite stark. There are only two ways to proceed. One is to accept that if present plans reach fruition, future controllers (in comparison to current ones) will have much less situational awareness, and therefore to plan now for reduced situational awareness in future air traffic control systems. Therefore, the safety and efficiency of future air traffic control systems must rely less on the controller's situational awareness. The other way to proceed is to preserve current levels of situational awareness, but this requires a positive policy and positive planning for it will not occur serendipitously. There do not seem to be any further practical options.

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406

Bisseret, A. (1970). Memoire operationelle et structure du travail. *Bulletin de Psychologie, 24*, 280-294.

Bisseret, A. (1971). Analysis of mental processes involved in air traffic control. *Ergonomics, 14*, 565-570.

Fitts, P. M. (1964) Perceptual-motor skill learning. In A.W. Melton (Ed.), *Categories of human learning*, (pp.243-285). New York: Academic Press.

Garland, D. J. (1991). Automated systems: The human factor. In J. A. Wise, V. D. Hopkin, & M. L. Smith (Eds.), *Automation and systems issues in air traffic control* (pp. 209-215). Berlin: Springer-Verlag.

Garland, D. J. & Stein, E. S. (1992). Air traffic controller memory: Implications for ATC tactical operations. *Proceedings of the 37th Air Traffic Control Association Conference* (pp. 132-139), Atlantic City, NJ.

Garland, D. J. , & Stein, E. S. (1993). Air traffic controller working memory: Considerations in ATC tactical operations. DOT/FAA Technical Report. DOT/FAACT-TN 93/37 Atlantic City International Airport, NJ.

Garland, D. J., Stein, E. S., Blanchard, J. W., & Wise, J. A. (1992). Situational awareness in the future air traffic control environment. *Proceedings of the 37th Air Traffic Control Association Conference* (pp.1032-1045), Atlantic City, NJ.

Garland, D. J., Stein, E. S., Blanchard, J. W., & Wise, J. A. (1993). Situational awareness in air traffic control: A critical yet neglected phenomenon. In D. J. Garland and J. A. Wise (Eds.), *Human Factors and Advanced Aviation Technologies*. Daytona Beach, FL: Embry Riddle Aeronautical University Press.

Garland, D. J. & Wise, J. A. (Eds.). (1993). *Human Factors and Advanced Aviation Technologies*. Daytona Beach, FL: Embry Riddle Aeronautical University Press.

Harper, R. R., Hughes, J. A., & Shapiro, D. Z. (1989). *The functionality of flight strips in ATC work*. Report to the U.K. Civil Aviation Authority. Lancaster, U.K.: Lancaster Sociotechnics.

Hollnagel, E., & Woods, D.D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies, 18*, 583-60.

Hopkin, V. D. (1988a). *Human factors aspects of the AERA 2 program*. Farnborough, Hampshire, U.K.: Royal Air Force Institute of Aviation Medicine.

Hopkin, V. D. (1988b). Air traffic control. In E.L. Wiener and D.C. Nagel (Eds.), *Human factors in aviation*. San Diego, CA.: Academic Press.

Hopkin, V. D. (1989). *Man-machine interface problems in designing air traffic control systems*. Proceeding IEEE, 77, 1634-1642.

Hopkin, V. D. (1991a). The impact of automation on air traffic control systems. In J. A. Wise, V. D. Hopkin, and M. L. Smith (Eds.), *Automation and systems issues in air traffic control*. (pp. 3-19). Berlin: Springer-Verlag.

Hopkin, V. D. (1991b) *Automated flight strip usage: Lessons from the functions of paper strips*. Paper presented at the conference on Challenges in Aviation Human Factors: The National Plan. Washington, DC: Federal Aviation Administration.

Hopkin, V. D. (1992). Human factors issues in air traffic control. *Human Factors Society Bulletin, 35*(6), 1-4.

Hutchins, E. (1986). Direct manipulation interface. In D. Norman & S. Draper (Eds.), *User centered system design: New perspectives in human-computer interaction*. Hillsdale, NJ: Erlbaum.

Intons-Peterson, M. J., & Fournier, J. (1986). External and internal memory aids: How often do we use them. *Journal of Experimental Psychology: General, 115,* 276-280.

Jackson, A. (1989). *The functionality of flight strips.* Report to the U. K. Civil Aviation Authority: Royal Signals and Radar Establishment. Farnborough, U.K.: U.K. Civil Aviation Authority.

Koriat, A., Ben-Zur, H., & Nussbaum, A. (1990). Encoding information for future action: Memory for to-be-performed versus memory for to-be-recalled tasks. *Memory & Cognition, 18,* 568-583.

Koriat, A., Ben-Zur, H., & Sheffer, D. (1988). Telling the same story twice: Output monitoring and age. *Journal of Memory & Language, 27,* 23-39.

Landis, D., Silver, C. A., Jones, J. M., and Messick, S. (1967). Level of proficiency and multidimensional viewpoints about problem similarity. *Journal of Applied Psychology, 51,* 216-222.

Langreth, R. (1993). Fail-safe skies: Coming in the 21st century? *Popular Science,1,* 76-86.

Lansdale, M. W., Simpson, M., & Stroud, R. M. (1990). A comparison of words and icons as external memory aids in an information retrieval task. *Behavior and Information Technology, 9,* 111-131.

Mogford, R. H. (1991). Mental models in air traffic control. In J. A. Wise, V. D. Hopkin, & M. L. Smith (Eds.), *Automation and systems issues in air traffic control* (pp. 235-242). Berlin: Springer-Verlag.

Moray, N. (1980). *Human information processing and supervisory control.* Technical Report. Cambridge, MA: M. I. T. Man-Machine Systems Laboratory.

Moray, N. (1988). Intelligent aids, mental models, and the theory of machines. In E. Hollnagel, G. Mancini, & D. D. Woods (Eds.), *Cognitive engineering in complex dynamic worlds* (pp. 165-175). London: Academic Press.

Narborough-Hall, C. S. (1987). Automation: Implications for knowledge retention as a function of operator control responsibility. *Proceedings of the Third Conference of the British Computer Society* (pp. 269-282). Cambridge, U.K.: Cambridge University Press.

Norman, D. A. (1986). Cognitive engineering. In D. A. Norman & S. W. Draper (Eds.), *User centered system design* (pp. 31-61). Hillsdale, NJ: Erlbaum.

Rasmussen, J. (1979). *On the structure of knowledge: A morphology of mental models in a man-machine systems context* (Report No. RISO-M-2192). Roskilde, Denmark: Riso National Laboratory.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human processing: I. Detection, search, and attention. *Psychological Review, 84,* 1-66.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 4,* 592-604.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352-373.

Vortac, O.U., Edwards, M.B., Fuller, D.K., & Manning, C.A. (1992a). *Automation and cognition in air traffic control: An empirical investigation.* Federal Aviation Administration Technical Report.

Vortac, O.U., Edwards, M.B., Jones, J.P., Manning, C.A., & Rotter, A.J. (1992b). *En route air traffic controller's use of flight progress strips: A graph-theoretic analysis.* Federal Aviation Administration Technical Report (DOT/FAA/AM-92/31).

Whitfield, D. (1979). A preliminary study of air traffic controller's picture. *Journal of the Canadian Air Traffic Controller's Association, 11,* 19-28.

Wickens, C. D. (1992). *Engineering psychology and human performance.* Glenview, IL: Scott, Foresman, and Company.

Wilson, J. R., & Rutherford, A. (1990). Mental models: Theory and application in human factors. *Human Factors, 31,* 617-634.

Wise, J. A., & Debons, A. (Eds.). (1987). *Information systems: Failure analysis.* Berlin: Springer-Verlag.

Wise, J. A., Hopkin, V. D., & Smith, M. L. (Eds.). (1991). *Automation and systems issues in air traffic control.* Berlin: Springer-Verlag.

Wise, J.A., Hopkin, V.D., & Stager, P. (Eds.). (1993). *Verification and validation of complex systems: Human factors issues.* Berlin: Springer-Verlag.

Mental Models and Situation Awareness in Air Traffic Control

Richard H. Mogford

CTA, Incorporated

Air traffic controllers report that they form a mental "picture" of aircraft in airspace to assist in their work. The relationship of the picture to the concepts of mental models and situation awareness is discussed. A proposed definition of situation awareness is, for any given system, the transient contents of awareness as structured and supported by an underlying mental model. Research on situation awareness in air traffic control is reviewed and it is suggested that insufficient work has been done to demonstrate its importance for successful task performance. An experiment with air traffic control trainees is described that relates their recall of basic aircraft data during a simulation to their scores in a final simulator exam. The results showed that students who remembered aircraft heading and altitude had better outcomes. It is suggested that trainees, in retaining information about aircraft altitude and direction of flight maintain a minimum set of data to help anticipate impending aircraft conflicts. It is recommended that further situation awareness research be conducted to determine its salient components for a given task.

Introduction

Maintenance of a safe and efficient flow of air traffic requires remote monitoring and control of aircraft from central locations. The air traffic controller is provided with various electronic devices, such as radar and radio, that collect and represent important information and allow the communication of instructions to aircraft. Using the data provided by these systems, controllers describe forming a mental "picture" of air traffic that assists with the conceptualization and prediction of aircraft movement. They state that maintenance of the picture is essential for effective air traffic control (ATC). In the human factors literature, such constructs are referred to as "mental models" or "situation awareness" (SA). This paper will explore the relationship between mental models and situation awareness and describe an experiment designed to demonstrate the importance of the air traffic controller's picture.

A mental model is a hypothetical construct that refers to an operator's learning and concepts about a system. Rouse and Morris (1986) defined mental models as "the mechanisms whereby humans are able to generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

system states" (p. 351). A mental model is an organized set of knowledge that has depth and stability over time. It is different from knowledge in general in that the term "model" suggests the formation of a conceptual analog of the external world in order to understand and predict system behavior.

In the case of ATC, there appear to be two components of the controller's mental model. The first is a "domain model" which encompasses to airspace, aircraft, and ATC procedures. The second factor is a "device model" which is an understanding of the electronic systems (including the computer-human interface) designed to support ATC. Both kinds of knowledge are essential if the air traffic controller is to accomplish the task of separating and guiding aircraft. This is analogous to the need to know some geography in addition to automobile operation in order to arrive successfully to a destination.

Rassmussen (1979) described mental models as having different levels of abstraction, including physical form, physical function, functional structure, abstract function, and functional meaning. Johnson-Laird (1983) distinguished between physical and conceptual mental models. This suggests that there may be a number of mental modeis of a system, ranging from the "concrete" (an analog, visual image) through various forms of increasing abstraction depicting functional relationships and operating rules, to a level subsuming the overall meaning and purpose of the system (Wilson and Rutherford, 1990). Although some type of model based on a visual image is probably involved in many air traffic control tasks, it is not be sufficient in itself for all aspects of the work. Conceptual, verbal, and numerical information must also be incorporated and applied as needed.

The term "situation awareness" is also one that is often used in the aerospace literature to discuss controller or pilot knowledge. How does this relate to mental models? One definition of SA is, "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley, 1988). It is evident from comparing this statement to the earlier quotation by Rouse and Morris (1986) that there may be a lack of distinction between mental models and SA. It would be useful to constrain the concept of SA to that which is implied in the term, i.e., the *contents* of awareness about a situation at any given moment. These contents might be data, meanings, or predictions but are not the same as the mechanisms which help *generate* this information. That is the function of the mental model.

In ATC, the mental model is the underlying knowledge that is the basis for SA or the picture. The controller's picture is defined by the underlying mental model and, in turn, supplies information to build and modify it. Sarter and Woods (1991) noted that mental models "...may be seen as the basis for adequate situation assessments which, in turn, result in flight-related knowledge that may eventually become part of the pilot's situation awareness. In other words, adequate mental models are one of the prerequisites for achieving situation awareness" (p. 49).

Based on the above discussion, it is possible to propose a simple model of human information processing about systems such as ATC as shown in Figure 1. The knowledge base or mental model is a complex array of information, some of which may not be consciously accessible. In ATC, the picture is a "holding area" where images and verbal data are maintained for ready use. "Momentary awareness" is the arena of ongoing conscious processes where there is no storage. Data from auditory or visual displays are momentarily perceived, held in SA (if they will be needed), and may update the mental model if there are long term implications. Information in the mental model influences and structures the data held in SA and directs attention. A relevant fact or rule may also emerge from this repository for storage in SA or for use in a momentary activity. As noted by the two arrows to the left of

the diagram, there is generally more consciousness as information moves up through the system although it is retained for a shorter period of time.

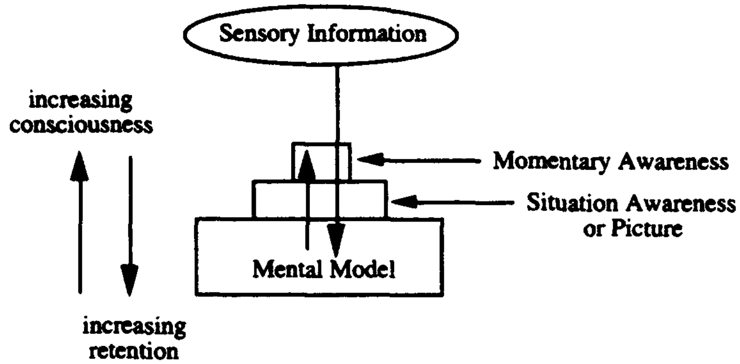


Figure 1. The relationship between the air traffic controller's situation awareness and mental model.

The concepts of the mental model and SA provide convenient ways of thinking about the operator's cognitive processes. However, apart from being interesting vehicles for theorizing about complex systems, there are practical issues to be addressed. How can these constructs be measured? Is there any evidence that the air traffic controller's ability to form a picture is an important skill? The following experiment focuses on the controller's SA and its relationship to job performance in ATC training.

In the human factors literature, there is a tacit assumption that good SA is important for almost any kind of skilled behavior. This seems reasonable and is supported by the statements of pilots, air traffic controllers, and others. However, it is also reasonable to question the relative importance of SA components and to investigate the relationship between SA degradation and operator error. In the ATC environment, it would be useful to determine which aspects of the controller's picture are critical for good performance.

Whitfield (1979) employed verbal data from controllers to study the picture and observed that most subjects reported it to be a three-dimensional, geographical representation which assisted with the understanding of the ATC situation. The identification of problems was also associated with the picture. It was sometimes described as a plan against which aircraft movement was compared and that helped in the detection of future conflicts. Whitfield (1979) concluoed that the picture assists the operator to feel in control, supplies information, and acts as a "supplementary display" in case of equipment failure.

One approach that has been employed to study human information processing in ATC is based on the operator's recall of radar screen information. Means, Mumaw, Roth, Schalger, McWilliams, Gagne, Rice, Rosenthal, and Heon (1988) found that controllers recalled enough information to identify 86% of the aircraft that had flown in a dynamic simulation exercise. Aircraft position and identifier were remembered with 84% and 24% accuracy, respectively. Although subjects were not requested to remember altitude and heading, when they spontaneously did so, it was with 79% and 80% accuracy, respectively. Bisseret (1970) found that highly qualified controllers had better recall for aircraft data than average controllers or trainees. Altitude and position information were reported most reliably.

Although these studies were not couched in terms of SA, many of the current procedures designed to assess SA are based on queries that require a subject to disclose information held in memory. One such technique is to freeze an ongoing simulation and probe the operator's situational knowledge.

The practice of freezing a simulation in order to assess SA has been employed in studies by Fraker (1988); Fraker (1989); Endsley (1990); and Marshak, Kuperman, Ramsey, and Wilson (1987). Sarter and Woods (1991) objected to this kind of approach because "these studies changed or even eliminated the phenomenon of interest, and therefore did not provide data about the natural character and occurrence of situation awareness" (p. 54). However, it might be argued that some of the embedded tasks listed by Sarter and Woods (1991) could also alter the phenomenon. These authors describe SA earlier in their paper by, "It [SA] refers to the accessibility of a comprehensive and coherent situation representation which is continuously being updated in accordance with the results of recurrent situation assessments" (p. 52). Given that SA contents should be accessible to the operator, interrupting an ongoing task to ask for information would seem to be at least one acceptable way of measuring it. Endsley (1990) demonstrated that freezing a flight simulation did not adversely affect performance, thus suggesting that this technique may not be overly intrusive.

Endsley (1989) defined three levels of SA. Level 1 involves the perception of situational elements or important facts in the environment. In Level 2, "Information Integration," the operator determines the meaning of the data. Level 3 SA is the "Projection of Future Status and Actions of Situational Elements." The experiment reported here does not aspire to address all three levels of a controller's SA, but attempts to determine if remembering basic hypothetically important Level 1 elements is related to successful control of aircraft as measured by a standard simulation-based evaluation. It could be argued that these basic data are the foundation for Levels 2 and 3 SA. Fraker (1989) conducted similar research which focused on factual data, such as aircraft position.

The following experiment was designed to assess the importance of the controller student's picture in the radar phase of ATC training. The hypothesis was that accurate recall of basic picture elements (Level 1 SA) would be an important predictor of ATC skill. Aircraft data of interest included altitude, speed, heading, identifier, and position.

Experiment 1

Method

Subjects. Two classes at similar stages of the radar phase of training at the Air Traffic Control School at the Transport Canada Training Institute (TCTI) in Cornwall, Ontario, Canada participated in the experiment. The first class had 20 students with mean age and years education of 26.4 and 14.9, respectively, and the second had 17 students with mean age and education of 28.1 and 15.1, respectively. There were 19 males and 1 female in the first class and 13 males and 4 females in the second class.

Materials and apparatus. The SA assessment procedure (or "picture assessment") employed a standard training exercise on the computer-based ATC simulation system at

TCTI. The simulator had multiple controller workstations so that it was possible to run up to six subjects during each experimental trial. An 11 inch by 17 inch map (on a cardboard backing) was used for student responses. The map was an enlarged photocopy of the Avalanche Sector (a fictitious sector of en route airspace used for training) showing fixes, airways, and the sector boundary, but no labels. It was used at the school for testing trainee knowledge of the airspace. Another set of maps was created for the simulation computer operators (or "pseudopilots") for the recording of the actual aircraft data.

The criterion measure of controller skill was an average of three final evaluation tests in the simulator. Each evaluation consisted of an over-the-shoulder rating of the student's ATC work by an instructor. Scores were comprised of four factors: aircraft separation, coordination, and planning errors as well as "affective factors" (which included separation visualization, planning priorities, board management/stripwriting, phraseology/communications, coordination/teamwork and confidence/comportment). The criterion was an average of the total scores of three final examination simulation runs.

Procedure. On the day of the picture assessment trials in the radar simulator, each class was assembled and given detailed instructions on the experimental procedure and response format. Subjects were informed that they would be involved in a standard, 45 minute simulation exercise and that it would be interrupted once at an undisclosed time. At the time of the freeze, subjects were asked to turn at least 90 degrees away from their radar screens (the screens were also dimmed) and flight data strips and were instructed to record aircraft identifier, position, altitude, speed, and heading in a specific format. Subjects were told that they should try to record as much aircraft data as they could as quickly as possible. Practice in the picture assessment procedure was provided prior to the experiment.

The pseudopilots were given identical maps and were asked carefully to note aircraft positions and other data from their displays when the simulation exercise was stopped. This allowed the experimenter to assess accuracy of recall. In addition, the instructors (there was one for each student) were asked to record the time each subject required to complete the recall task.

The two classes were each split into four groups for testing. Prior to the experimental trial a map was given to each subject and simulator pilot and they were reminded of the procedure. After about 20 minutes (when there were eight active aircraft) the exercise was interrupted and subjects used their maps to record information about the aircraft they had been controlling. Pseudopilots recorded the actual data on each aircraft. After these tasks were completed, the exercise was restarted and completed.

Accuracy of recall was scored using a "ballpark" method developed in consultation with ATC instructors. The criteria were as follows: identifier, two or more letters/numbers correct *and* in order; position, within 15 nautical miles; altitude, plus/minus 1000 feet; speed, plus/minus 50 Knots; and heading, plus/minus 10 degrees.

Results

The mean accuracy of recall of aircraft information during picture assessment was as follows: position, 86%; heading, 82%; altitude, 73%; identifier, 55%; and speed, 53%. (Some of the position information was contaminated by measurement problems and was removed from the analysis. However, sixty-nine percent of the position data were unaffected and there was no problem with the other variables.)

Stepwise multiple regression was applied to investigate the relationship of the picture assessment variables to the measure of success in radar training. Table 1 shows the results. The regression model accounted 38% of the variance in the outcome measure (based on adjusted R^2) and the test of the regression was significant. Three of the seven picture assessment measures (speed, altitude, and heading) earned significant regression coefficients ($p < .05$). Not included were position, identifier, number of aircraft recalled, and rate of recall. (One subject dropped out of the study. Two of the remaining 36 cases were excluded from the data set after regression diagnostics identified them as outliers. These subjects scores were atypical compared to the other subjects in the regression model.)

Table 1. Average Radar Simulation Performance Regressed Over Predictor Variables

MULTIPLE R	0.66
MULTIPLE R^2	0.44
ADJUSTED R^2	0.38
SE OF ESTIMATE	4.36

ANALYSIS OF VARIANCE	SS	df	MS	F RATIO
REGRESSION	444.76	3	148.25	7.79
RESIDUAL	570.81	30	19.03	p=.0005

VARIABLE	b	SE	b	t	p (2 TAIL)
HEADING	0.13	0.04	0.50	3.27	0.003
ALTITUDE	0.12	0.04	0.49	3.11	0.004
SPEED	-0.06	0.02	-0.44	-2.62	0.014
CONSTANT	64.59				

Discussion

Three of the picture assessment variables were included as significant predictors of ATC competence, thus partially supporting the hypothesis that good Level 1 SA forms a basis for competent air traffic control in a training environment. The regression model indicated that good mental representation of the direction of travel and altitude of aircraft targets on the radar display were important. However, it seemed that the recall of speed was counterproductive. Investigation of the intercorrelations of the variables, however, suggests that speed played the role of a "suppressor variable" in the regression equation.

A suppressor variable is one that has little or no correlation with the criterion (or Y variable) but is correlated with one of the predictor variables. This correlation adds irrelevant

variance to the predictor and reduces its relationship with Y. "The inclusion of the suppressor variable in the analysis increases the partial correlation because it serves to suppress, or control for, irrelevant variance, that is, variance that is shared with the predictor and not with the criterion, thereby ridding the analysis of irrelevant variation, or noise" (Pedhauser, 1982, p. 104). The aircraft speed variable fits this description in that its correlation with the criterion measure (simulator score) was $r = 0$ but its correlations with heading and altitude accuracy were $r = .42$ and $.48$, respectively. Therefore, when interpreting these results, the negative regression weight for speed in Table 1 can be ignored given that the presence of this variable only serves as a catalyst to improve the correlation of heading and altitude with simulation score.

Remembering aircraft speed may have something in common with recalling heading and altitude; it could draw upon the same abilities that support Level 1 SA. The lack of any correlation between speed and the criterion measure suggests that retention of speed information in controller SA is not practically useful when ATC skills must be applied. Including it in the regression equation may remove this common factor and help demonstrate the utility of mentally representing heading and altitude facts when controlling aircraft.

The number of aircraft recalled, their identifiers, and positions also did not have much weight in the equation. Identifier did not emerge as an accurately remembered piece of aircraft data in other ATC studies (Means, et al. 1988; Bisseret, 1970). Judging by the results of these same studies, however, it is surprising that position was not more important. Its simple correlation with the criterion measure was low ($r = .21$), and it was not represented in the regression equation. The loss of data due to measurement problems may have affected its performance.

Speed (apart from its roles as a suppressor variable) also did not appear to be critical. Sperandio (1978) observed in his research on workload in ATC that this variable which was poorly represented by controllers as workload increased. Although aircraft speed can be important, especially in relation to other aircraft, it does not play as critical a role in maintaining separation as altitude.

Conclusions

The results of this experiment show that, on the average, ATC trainees were able to report aircraft position, heading, and altitude with at least 73% accuracy during a freeze in a real-time ATC simulation. They were also able to recall identifier and speed for about one-half of the aircraft. Thus there is evidence that these students were able to maintain good SA of a number of aircraft attributes.

However, when multiple regression was used to identify which of these variables was important for success in a final simulator exam, it was found that those ATC trainees with good SA of aircraft altitude and heading achieved the highest scores. This suggests that certain Level 1 SA aircraft data are more critical than others in this setting. Although this experiment was limited in scope in that it did not measure Levels 2 and 3 SA, it begins to address a critical issue in SA research, that of empirically demonstrating the importance of SA in supporting the performance of specific tasks.

Although it might be expected that all aircraft information is critical for adequate air traffic controller SA, this experiment demonstrates that some elements (such as aircraft altitude and heading) may play a key role while others (such speed, position, and identifier) may not be as important as expected. Perhaps it is sufficient, in order to anticipate aircraft conflicts, for student controllers only to maintain SA regarding the directions and altitudes of targets. An analysis of conflict detection in ATC suggests that the following sequence of tests are required (all four must be true for a separation conflict to be identified):

1. Are the aircraft at the same altitude?

2. If so, are they on converging headings?

3. If so, do their relative positions indicate impending conflict?

4. If so, do their relative speeds indicate impending conflict?

Successful air traffic control students may learn to maintain the basic aircraft facts sufficient to make the first two tests on the above list. Given the heavy cognitive demands of ATC, these trainees reduce SA requirements to a minimum for the sake of efficiency. Operators of complex systems may not, therefore, work to retain all of the available information that a task analysis might deem important. They could instead develop a strategy to retain critical SA information that will improve their performance of high priority tasks. There may then be three kinds of data in the ATC environment: that which must be remembered and updated, that which can be searched for when needed and then forgotten, and that which can be ignored. Only the first type of data is retained in SA. Measuring mental models may never be practically possible if it can be assumed, as suggested in the introduction of this paper, that they are complex knowledge structures that may, at least in part, be inaccessible to conscious exploration. SA, being a dynamic and transient set of information that is supported by underlying mental models, is more amenable to various kinds of assessment. However, there appears to be a temptation in SA research to make many assumptions about its importance. Only further investigation that links SA components at each of the three levels to task performance can demonstrate its usefulness and show us more about the creativity and intelligence inherent in operator strategies.

Acknowledgments

This research was part of a doctoral dissertation completed at Carleton University, Ottawa, Canada. The assistance of Brian W. Tansley and Transport Canada Air Traffic Services was greatly appreciated.

References

Bisseret, A. (1970). Mémoire opérationelle et structure du travail. *Bulletin de Psychologie, 24*, 280-294.

Endsley, M. R. (1990). *Situation awareness in dynamic human decision making: Theory and Measurement* (NOR DOC 90-49). Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (1988). *A construct and its measurement: The functioning and evaluation of pilot situation awareness* (NOR DOC 88-30). Hawthorne, CA: Northrop Corporation.

Fraker, M. L. (1988). A theory of situation assessment: Implications for measuring situation awareness. In *Proceedings of the Human Factors Society 52nd Annual Meeting* (pp. 102-106). Santa Monica, CA: Human Factors Society.

Fraker, M. L. (1989). Attention allocation in situation awareness. In *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 1396-1399). Santa Monica, CA: Human Factors Society.

Johnson-Laird, P. N. (1983). *Mental Models*. Cambridge, England: Cambridge University Press.

Marshak, W. P., Kuperman, G., Ramsey, E. G., & Wilson, D. (1987) Situational Awareness in map displays. In *Proceedings of the Human Factors Society 31st Annual Meeting* (pp. 533-538). Santa Monica, CA: Human Factors Society.

Means, B., Mumaw, R., Roth, C., Schalger, M., McWilliams, E., Gagne, E., Rice, V., Rosenthal, D., & Heon, S. (1988). *ATC training analysis study: Design of the next-generation ATC training system* (OPM Work Order No. 342-036). Richmond, VA: HumRRO International.

Pedhauser, E. J. (1985). *Multiple Regression in Behavioral Research.* New York, NY: Holt, Rinehart, and Winston.

Rasmussen, J. (1979). *On the structure of knowledge: A morphology of mental models in a man-machine system context* (Report No. RISO-M-2192). Roskilde, Denmark: Riso National Laboratory.

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin, 100*, 349-363.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *International Journal of Aviation Psychology, 1*, 4[illegible]-57.

Sperandio, J. (1978). The regulation of working methods as a function of work-load among air traffic controllers. *Ergonomics, 21*, 195-202.

Whitfield, D. (1979). A preliminary study of the air traffic controller's picture. *CATCA Journal, 11*, 19-28.

Wilson, J. R., & Rutherford, A. (1990). Mental Models: Theory and application in human factors. *Human Factors, 31*, 617-634.

The Role of Dynamic Memory in Air Traffic Controllers' Situation Awareness

Esa Rantanen

Embry-Riddle Aeronautical University

Introduction

Although the term situation awareness is fairly new, air traffic controllers have always been familiar with the concept. Controllers refer to their situation awareness as the "picture", either "having the picture" or "losing the picture". The current definitions of situation awareness (e.g., Endsley, cited in Garland, Phillips, Tilden, and Wise, 1991) quite accurately describe the demands of air traffic control (ATC) as well, and make the controller's picture synonymous with situation awareness. In the following text, the terms "picture" and "situation awareness" are used interchangeably.

The air traffic controller's memory is very important in ATC in general, but especially in the maintenance of the controller's picture. Because of the nature of the job and the environment in which controllers work, their dynamic memory in particular has a central role in updating the picture. There is, however, very little research done on the relationships between the controllers picture, dynamic memory, and workload, all of which are fundamental concepts of ATC. This paper will attempt to suggest some of these relationships and raise questions about the role of the dynamic memory in the air traffic controller's situation awareness.

The Controller's Picture

The air traffic controller's picture is a mental model of the airspace architecture, layout of the runways at airports, rules and standard procedures regulating the conduct of flights, and positions, flight data, and performance characteristics of the aircraft operating within this system. Also included in the picture are numerous other factors relevant to the traffic situation, such as the weather, operational status of navigation aids and the ATC equipment, staffing, and sectorization within the facility, and possible irregularities within these.

Endsley's (cited in Garland, Phillips, Tilden, and Wise, 1991) definition of situation awareness as "the perception of elements in the environment within a volume of time and

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

space, their comprehension and meaning, and the projection of their states in the near future" summarizes the essence of ATC. Controllers have to keep track of several aircraft simultaneously, sort the available information according to its importance, and prioritize their actions within constrained space and time. Furthermore, the situations change continuously, sometimes at a very rapid rate, making it difficult to judge the ultimate importance of the information.

However, it is important to note that the bits of information controllers receive through displays and communication channels are hardly new to them. In fact, most information is expected and it fits in the controller's mental model--or picture--immediately upon receipt, reducing processing requirements to mere acknowledgment of the data. The secret of the sometimes astonishing mental performance of air traffic controllers thus lies in their picture and in the relevance of the information coming into that picture.

Controllers seem to work on two mental levels simultaneously, with each level facilitating the functions of the other. Sperandio (1978) identifies these levels as a process of decision and a process of action. Preplanning, or the projection of aircraft trajectories into the future and subsequent decision-making, is the basis of controller performance. When the situations have been analyzed and planned ahead of time, monitoring the events actually taking place and acting upon them requires little effort and attentional resources. The current situation is already familiar, freeing resources for further planning tasks.

Another important factor is the familiarity with the supporting structures, i.e., the airspace architecture, rules and standard procedures, and flight plan data. This kind of knowledge base immediately puts the incoming information in the right context, facilitating rapid processing, effective chunking, and good situation awareness. Controllers also build their own structures within the existing framework when working traffic. They create patterns that result in smooth and conflict-free traffic flows according to each situation. All this will become part of the controller's picture.

The patterns along which traffic is controlled vary extensively depending on many factors, e.g., weather and mix of traffic. However, these patterns are usually modified from a few relatively fixed patterns, which are determined by general directions of traffic flow, airway and navigation aid structures, and airport layouts. Experienced controllers are very familiar with these patterns, and modify them only as necessitated by a given situation. This suggests a definite skilled memory effect associated with air traffic controller performance. Garland and Stein (1991) observe that controllers do not necessarily process the information as thoroughly as would appear from their decisions. As in case of chess masters (Chase, 1986), controllers may need very little information to recall a similar situation or pattern from their long-term memory and then use this information as a basis for their decisions.

As has been discussed before, controllers rely heavily on their experience and knowledge base in the formation of their picture. The picture, being a conscious part of the controller's mental model, exists in the working memory (Mogford, 1991). However, the role of dynamic memory as a manager of information flow between working and long-term memories becomes critical in the formation and maintenance of the picture.

Dynamic Memory Defined

Dynamic memory appears to be best defined in terms of circumstances. Situations where the flow of information is continuous, the information changes or is updated frequently, and a great number of variables are included, place considerable demands on a person's dynamic

memory (Moray, 1986; Wickens, 1992). A feature of the dynamic memory is the need for active management of the information held in it, including the prompt discarding of old and unimportant information (Hopkin, 1980).

Yntema (1963) performed a series of experiments on memory capacity under the above-mentioned circumstances. Subjects were asked to keep track of a large number of objects, each varying along certain attributes, which in turn would have a certain value. The results of these experiments suggest that the capacity of the running, or dynamic, memory is even smaller than that of the static working memory: the subjects would make mistakes while keeping track of only two or three things at once. Further, Yntema observed that performance was not much improved even when the variables followed a certain regular and predictable pattern. However, performance was improved when the number of objects was reduced, even when the number of attributes remained the same.

Yntema's experiment is quite analogous with the tasks of air traffic controllers. In fact, the task in the experiment was modeled after an ATC situation (Yntema, 1963). However, the subjects in the actual experiment were tasked with remembering completely meaningless information, such as unrelated letters, shapes, colors, animal and food names, and the like, instead of flight data relevant to a realistic ATC situation. This may have affected the results and conveys an overly pessimistic view of human capabilities.

Moray (1986) defines dynamic memory as keeping track of a great deal of information arriving in a continuous stream without a definite interval for recall. This definition also serves as an apt description of an air traffic control task. Consistent with Yntema's experiments, he concludes that observers viewing a time series are not capable of holding more than three items in their dynamic memory. Also Wickens (1992) identifies a running memory as a memory handling random stimuli, where a different response is required for each stimulus or series of stimuli some time after they have occurred. In his text, running memory is synonymous with dynamic memory.

Air traffic controllers, however, seem to defy the paradigm of limited working and dynamic memory capacities. Moray (1986)--although generally agreeing with Yntema's conclusion that the dynamic memory's capacity is only three items--notes that these results may have been due to the fact that the items presented to the subjects were random and meaningless to them. He had observed a significantly higher capacity of the dynamic memory among air traffic controllers and suggests that this may be due to the fact that by actively handling the flights, controllers in a sense generate the information to be kept in the dynamic memory. Based on Megaw and Richardson's (1979) experiment on visual scanning strategies, Moray also suggests viewing the gathering of information as "a cumulative process, but whose outcome was the convolution of data acquisition function and a forgetting function" (p. 40-28).

Garland and Stein (1991) further emphasize the importance of the meaningfulness of information. Air traffic controllers, for example, may be able to enhance their dynamic memory capacities simply by chunking information more effectively. Chunking is useful in two different ways: first, it helps to maintain information in the working memory longer, and second, it facilitates the transfer of information to the long-term memory for more permanent storage (Wickens, 1992). If information is meaningful, it can be encoded more effectively into larger chunks, which in turn allows more effective storage and retrieval strategies between the working and long term memories. Dynamic memory appears to have a key role in both of these processes.

Hopkin (1980) also addresses the importance of forgetting as a part of managing the working memory. Given the highly dynamic environment of ATC and the rapid pace of

information update, old information must be effectively dumped from the memory to make room for new, more critical information. This kind of active management appears to be characteristic of dynamic memory.

In general, the available literature consistently stresses the importance of the meaningfulness of information. This implies that dynamic memory might have a very important and central role as a facilitator of information transfer between working and long-term memories (Figure 1). Further, this emphasizes the importance of both experience, i.e., the knowledge base in the long-term memory, and the quality of information received. Quality of information in this context means that it can be rapidly and easily coded to match the information in the long-term memory.

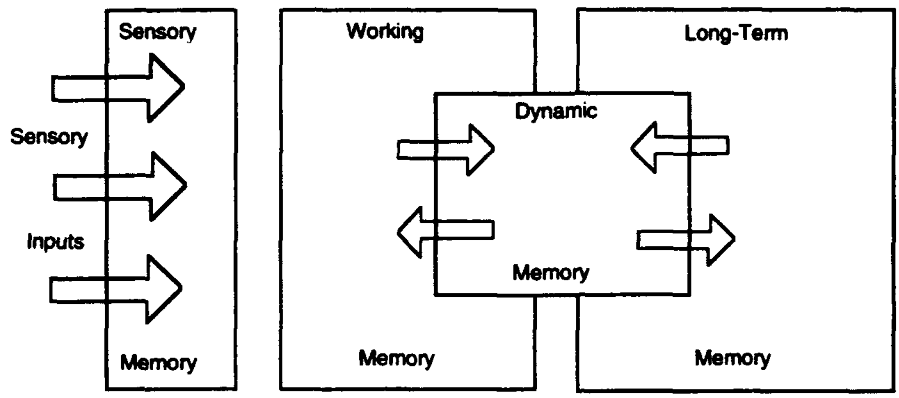


Figure 1.

The literature and the proposed role of dynamic memory raise a question of the driving forces behind it. Sufficient workload and level of activity appears to have a definite effect on the performance of the dynamic memory. Also the multiple resources theory (Mane & Wickens, 1986) offers some insights in understanding the functions of the dynamic memory. The work of Mane and Wickens (1986) concerns training situations, but their findings are applicable to performance in other tasks as well. They found that workload during learning and the difficulty of the task have severe implications on the learning performance. If the task is difficult, more resources will be allocated to its performance and it will be learned better (Mane & Wickens, 1986). However, this is true only when the difficulty stems from the task to be learned. If the trainee has to perform other, secondary tasks not benefiting learning, resources will be deployed away from the learning situation resulting in a negative impact on the learning performance. This suggests the importance of attention in the performance of the dynamic memory and maintenance of the controller's picture.

Workload, Dynamic Memory, and Situation Awareness

Many pilots flying most modern aircraft equipped with "glass cockpits" have reported deterioration of their flying skills as a result of increased automation. Increased automation of ATC may result in similar phenomena among controllers. Without direct involvement

with traffic, controllers--especially younger ones--may never be able to develop the skilled database in their long term memories necessary for the efficiency of the working memory. Following Craik's and Lockhart's (1972) theory of levels of processing, even experienced controllers may find it difficult to utilize their dynamic memory capacities fully if the level of direct manipulation and interaction, and therefore the level of information processing, is reduced.

A significant amount of incidents in ATC happen during times of low traffic levels and low workload. This is hardly surprising, since it is a well known fact that human reliability deteriorates rapidly in tasks requiring continuous maintenance of attention over long periods of time without much overt action (Hopkin, 1982). This further suggests an identifiable optimum workload for the best dynamic memory performance. It can be speculated that this workload can be fairly high, resulting in concentrated attention on the task at hand and minimizing the susceptibility to external distractions.

Sperandio (1971, 1978) has observed controllers adapting to increased workload by changing their operating strategies. As the amount of traffic under their responsibility increases, controllers become selective of the information they process and deal with only the most relevant variables associated with each individual flight. Furthermore, they begin to treat individual aircraft as links in a chain whose characteristics remain rather stable. It can be argued that this kind of more economical working strategy also results in better situation awareness. Controllers handling strings of aircraft rather than individual flights see the "big picture" and are able to better predict the effect of an individual flight on the traffic flow. This kind of highly structured handling of traffic also typically results in fewer conflict points to be monitored.

There are several common control techniques that support the controllers' information processing and memory, and which result in improved situation awareness. As stated in Sperandio's (1971, 1978) findings, controllers do not only "chunk" information given to them, but literally chunk aircraft they have to monitor. One of the most common methods used by approach controllers is to arrange inbound aircraft on downwind, at the same altitude, at the same speed, and at sufficient distance apart in trail. Thus, instead of keeping track of many individual aircraft, the controller has to keep track of a single string of aircraft, a string which behaves exactly as planned. When it comes time to turn the aircraft on final, the controller simply takes one aircraft at a time from downwind as it approaches the turning point and gives it a heading for localizer interception and approach clearance. Anderson (1991) claims that with these kind of techniques there is practically no limit to the number of aircraft a controller can handle. Although the above example is perhaps overly simple and applies to only a few situations, it nevertheless illustrates some ways controllers have developed to counter memory overloads. It also illustrates the importance of experience and patterns stored in the long-term memory, and the role of the dynamic memory in tapping this data base while constructing the picture in the working memory.

Conclusions

There appears to be a definite interrelationship between the air traffic controller's situation awareness, dynamic memory, and workload. The available literature quite consistently

stresses the importance of the meaningfulness of information for efficient processing. This implies that the dynamic memory might have a very important and central role as a facilitator of information transfer between the working and long-term memories. Further, this emphasizes the importance of both the experience (i.e., the knowledge base in the long-term memory) and the quality of information received. Quality of information in this context means that it can be easily and rapidly coded to match the information in the long-term memory.

Given these assumptions and the active nature of the dynamic memory, it seems natural to conclude that active involvement with the task at hand is essential in supporting the dynamic memory. When air traffic controllers handle traffic, for example, they have an overall picture of the situation in which even the smallest bits of information are within a context, and therefore meaningful. One of the greatest dangers of automation is that the operator is removed so far from the situation that this contextuality is lost.

References

Anderson, R. (1990). Vector and speed control techniques at Honolulu. *ATCA Journal* (pp. 38-40). April-June 1990.

Chase, W. G. (1986). *Visual information processing*. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Volume II, Cognitive processes and performance* (pp.28-53 -28-57). New York: Wiley Interscience.

Craik, F. I. M. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.

Garland, D. J., Phillips, J. N., Tilden, D. S., & Wise, J. A. (1991). *Theoretical underpinnings of situational awareness: Towards an objective measure*. (Tech. Rep. No. CAAR-15498-91-1). Daytona Beach: Embry-Riddle Aeronautical University, Center for Aviation/Aerospace Research.

Garland, D. J., Stein, E. S., Wise, J. A., & Blanchard, J. W. (1991). *Situational Awareness in air traffic control: A critical yet neglected phenomenon*. In D.J. Garland & J. A. Wise (Eds.), *Human Factors and Advanced Aviation Technologies*. Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Hopkin, V. D. (1980). The measurement of the air traffic controller. *Human Factors*, 22(5), 547-560

Hopkin, V. D. (1982). *Human factors in air traffic control.* Paris: NATO, AGARDograph no. 275.

Mane, A. & Wickens, C. D. (1986). The effects of task difficulty and workload on training. *Proceedings of the Human Factors Society 30th Annual Meeting*, 1124-1127.

Megaw, E. D. & Richardson, J. (1979). Target uncertainty and visual scanning strategies. *Human Factors, 21*(3), 303-315.

Mogford, R. H. (1990). *The characteristics and importance of the air traffic controller's mental model*. Unpublished doctoral dissertation, Carleton University, Ottawa, Ontario.

Moray, N. (1986). Monitoring behavior and supervisory control. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Volume II, Cognitive processes and performance* (pp. 40-1 - 40-51). New York: Wiley Interscience.

Sperandio, J. C. (1971). Variation of operator's strategies and regulating effects on workload. *Ergonomics, 14*, 511-517.

Sperandio, J. C. (1978). The regulation of working methods as a function of workload among air traffic controllers. *Ergonomics*, *21*, 195-200.

Stein, E. S. & Garland, D. J. (1991). *Air traffic controller working memory: Considerations in air traffic control tactical operations. DOT/FAA CT-TN93/37. FAA Technical Center: Atlantic City International Airport, NJ.*

Wickens, C. D. (1992). *Engineering psychology and human performance* (pp. 211- 257). New York: Harper Collins Publishers.

Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors, 6,* 7-17.

SATORI: Situation Assessment Through The Re-Creation Of Incidents

Mark D. Rodgers [1] & Duane A. Duke [2]

[1] Civil Aeromedical Institute, Federal Aviation Administration
[2] Office of Information Services, Federal Aviation Administration

SATORI: Situation Assessment Through Re-Creation Of Incidents

Introduction

A system has been developed that graphically re-creates the radar data recorded at En Route air traffic control (ATC) facilities. This data is sent to the controller scope, called the Plan View Display (PVD), and the Continuous Readout Update Display (CRD). The re-creation synchronizes the graphic display of this data with tapes containing the associated verbal interactions between pilots and the controller. This system is called Situation Assessment Through Re-creation of Incidents (SATORI). SATORI is a Japanese word that translates to English as "insight". In Zen Buddhism it refers to a "state of enlightenment". It is hoped that through the use of the SATORI system it will be possible to gain a better understanding of the interaction between the various elements of displayed information, verbal interactions, and the control actions taken by air traffic control specialists (ATCSs).

Description of ATC System

This development uses data obtained from Air Route Traffic Control Centers (ARTCCs), also called En Route facilities. ATCSs at those facilities primarily handle aircraft traveling between the terminal facilities across the nation. Each facility records PVD and CRD data associated with the airspace under its control on a System Analysis Report (SAR) tape. The SAR tape contains all of the recorded dynamic display information about the National Airspace System (NAS), including the weather and the aircraft traversing it. Verbal interactions between pilots and controllers are recorded on a multi-channel tape unit at each facility.

ATCSs are required to maintain certain separation minima between aircraft under their control. Standards for separation minima are described in the Air Traffic Control (ATC) Handbook (7110.65G, and supplemental instructions). While there is considerable

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

complexity in those standards, at flight levels between 29,000 and 45,000 feet, Air Traffic Control Specialists (ATCSs) at En Route facilities are required to maintain either 2,000 feet vertical separation or 5 miles horizontal separation between aircraft. At flight levels below 29,000 feet with aircraft under instrument flight rules (IFR) conditions, ATCSs are required to maintain either 1,000 feet vertical separation or 5 miles horizontal separation. An operational error (OE) takes place when an ATCS allows less than the prescribed minimum separation distances between aircraft (or an aircraft and an obstruction).

Analysis of Operational Errors

Currently, the FAA Office of Air Traffic Systems Effectiveness requires an investigation into each OE. This investigation involves determining the circumstances in which the OE occurred and the causal factors associated with the error. Initially, a preliminary investigation report (FAA 7210-2) is filed in which possible causal factors are identified and a final report (FAA 7210-3) filed shortly thereafter. A project related to the development of SATORI and sponsored by the FAA Office of Air Traffic Systems Effectiveness involves studying the tasks of an En Route ATCS associated with the commission of an OE. One way to analyze OEs would be to identify which tasks were omitted or were performed incorrectly. This may facilitate identification of training needs or system deficiencies. In order for the tasks associated with the commission of an OE to be identified, it must become possible for one to review the dynamics of the situation in which the irregularity occurred.

Prior to the development of SATORI, it was not possible for the Quality Assurance (QA) team investigating errors to review how the control situation was seen by the ATCS as the OE occurred. That is, the dynamics (the interaction between control actions and displayed data) of the situation were unavailable for review, not only by the QA team investigating the irregularity but also by the controller who committed the error. This limited not only the extent to which a determination could be made of the tasks involved in an error, but also the effects of the dynamic situation on ATCS situation awareness.

In addition, the only means by which a graphical representation of an En Route OE could be achieved was to obtain a printout of the National Track Analysis Program (NTAP) or to have a simulation built at the ARTCC or FAA Technical Center using dynamic simulation (DYSIM) equipment. NTAP processes NAS data recorded on the SAR tape and provides a plot of aircraft tracks and altitude information as output using a line printer. NTAP is limited to the display of about four aircraft; however some of the data are lost when all four aircraft are displayed since it was designed not to overwrite information already printed. When only two aircraft are displayed, the information loss due to printing is minimal. The DYSIM simulation built at the FAA Technical Center or ARTCC is not as timely or accurate as NTAP. Each piece of data and its associated track must be hand entered to build a simulation using DYSIM equipment. The simulation would only be as accurate as the data used to create it, and since this is an extremely labor intensive process, typically not all of the data points are used.

Without the ability to review an error with the involved controller, the dynamics between the situation and control actions taken, as well as the task elements involved in the OE, remain relatively obscure. The purpose of SATORI is to display the ATC situation dynamics so that a more definitive determination of the factors involved in OEs becomes possible. SATORI utilizes a multi-media graphics workstation which has the capability for developing

a library of OEs and an OE performance and taskload database. A diagram of the SATORI data processing flow is provided in Figure 1.

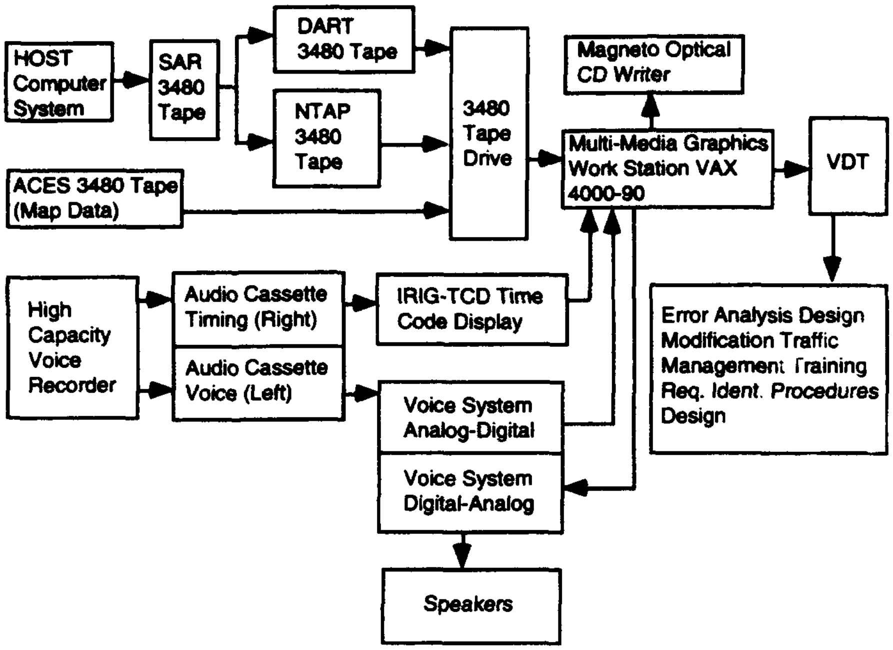


Figure 1. SATORI Data Process Flow Chart

SATORI Development

The following is a discussion of the SATORI development process. First, current system capabilities are described along with those under development. A brief note is included for issues yet to be resolved concerning capabilities under development. A detailed discussion of the National Airspace System Program (NASP) data files required for the development is provided when relevant. Also included in this section is a discussion of the required hardware developments and graphics software requirements for SATORI. The second section discusses the potential applications of SATORI. A concluding section discusses the future developments and potential directions of the SATORI project.

Graphical Re-creation of Airspace

For the SATORI development, the SAR tape is edited using the Data Analysis and Reduction Tool System (DART) and NTAP to obtain only those files required to provide the information displayed on the PVD and CRD. Several data files are required to obtain the

necessary information for re-creating the information presented on the PVD. Aircraft position information is recorded in three separate files, one for each of the three symbols that could potentially be presented on the PVD for any given aircraft. These three symbols are the aircraft position symbol, beacon target, and primary target.

The tracks of the aircraft position symbols as they move through the facility airspace are recorded on the SAR tape for aircraft under positive control and are available using DART; the file is called TRACK. Position symbols are updated every 6 seconds and represent predicted aircraft positions based on the current aircraft track. This file also specifies the information contained in the data block and the direction of its leader line. DART does not allow for the extraction of aircraft primary radar or beacon targets, which are also recorded on the SAR tape. These data files are extracted using NTAP. Primary targets represent the position of aircraft based on radar data. Primary targets are required for the display of aircraft not under positive control (i.e., uncorrelated targets) in the sector being re-created. Primary targets are updated on the PVD every 10, 11, or 12 seconds depending on the aircraft location relative to the radar site and represent actual aircraft location to the degree of accuracy associated with the system. However, since the data associated with primary targets include only a symbol type and location without any identifier, representing them with SATORI may be inaccurate because it is not possible to tell how long a primary target symbol should be displayed (i.e., 10, 11, or 12 seconds). Several solutions to this problem are currently being considered, however further work is required.

Beacon targets represent aircraft locations generated by transponder-equipped aircraft and are used by controllers to provide separation between aircraft. Beacon targets for all transmitting transponder-equipped aircraft are displayed with SATORI. Currently, the SATORI system allows for review of all controlled aircraft within any sector of airspace for any given period of time that was recorded on a SAR tape.

SATORI software overlays the NAS SAR data on the appropriate sector map using map data from the Adaptation Control Environmental System (ACES) database. This database is used to generate the various map configurations displayed on the PVD. The ACES map files contain map data for each sector within a given facility for all ARTCC facilities. These maps contain airway, navaid, and airport information, and other types of data used in the graphical representation of the airspace of the NAS. The file of particular interest in this database is the GMLMAP file. This file contains the logical map records for the center of interest.

Software has been developed that allows for any sector map to be edited from the ACES database for any ARTCC. This software allows for the selection of four maps which can be displayed together or separately, as follows: (1) sector boundary; (2) airways; (3) low sector boundaries below a high altitude sector or the high sector boundaries above a low altitude sector; and (4) any other available map data such as military operations areas (MOAs) and restricted areas.

Audio Data Synchronization

SATORI provides for the synchronization of the audio and video portions of the replay. The audio and video displays have the capability to start at any time, pause, stop, and reset to the earliest recorded time. Synchronization routines have been written that keep the audio and video displays synchronized to within 1 second (the audio time signals are only accurate to the second).

Audio data are officially recorded on two channels, with all of the voice communications between pilots and the controller for a given sector on the left channel and the Inter-Range Instrument Group (IRIG) timing signals on the right channel. These signals or reference pulses are amplitude-modulated time codes involving a 600 Hz. carrier signal (FAA modified IRIG-E). The signals are read by a Time Code Display (TCD) unit. Datum model 9700 TCD is equipped with RS-232 binary output. The Datum 9700 output provides the data to synchronize the voice channel to the video presentation. SATORI can also incorporate interphone communications into the re-creation.

The voice channel is first digitized, synchronized through timing routines to the video presentation, and then converted back to an analog signal for replay. The analog-to-digital, digital-to-analog data acquisition system is made by Gradient Technology.

PVD Emulation Subroutines

The analog switch display settings of the PVD are not recorded; however, subroutines have been written for SATORI that allow the display to be set up with the settings reported to have been used by a given controller. These include the vector velocity line, leader line, history, display center, range, and brightness. The vector velocity line can be set for one through five minutes, and the leader line length has five settings. The center of the display can be set anywhere in the sector map area selected from the ACES database. Range is selectable from 6 to 400 miles; however, data are edited from the SAR according to the PVD device number and therefore might not be available for display at all ranges. Brightness is adjustable only as a function of the workstation's monitor controls. SATORI also allows for the display of the J-rings a controller selected for display.

Digital settings on the PVD, such as altitude filter selections, display of weather, alpha-numeric keypad (ANK) entries, use of quick action keys (QAK), and any changes made to the digital settings, are recorded on the SAR tape. Using DART to obtain the LOG file from the SAR tape, it is possible to identify specifically what those settings and changes were and when they occurred. The SATORI system allows for display of the LOG file data, which are the digitally recorded Host Computer System (HCS)/ATCS interactions. Most of these data are displayed on the CRD display located next to the PVD display. The CRD displays all data entered using the alpha-numeric keypad and quick action keys (QAK). Both the radar and data positions can make QAK and ANK entries which are displayed on separate CRDs. SATORI displays the CRD data from both positions simultaneously on separate CRD displays. HCS response to requests for route readouts and arrival/departure lists is not currently part of this development.

In addition to the above, SATORI has the capability to display the high and low weather intensity that was displayed on a given PVD. This should be particularly helpful in reviewing situations requiring pilot deviations from typical routes when navigating around weather. Both heavy and light weather symbols are available for review with data obtained using NTAP. Conflict alerts for a given sector will be available for display with data obtained using DART.

Hardware and Software Systems

All software routines written for SATORI use Open Systems Foundation (OSF) technology. The OSF standard recommends the use of an operating system with Posix compliance, ANSI C programming language, the X-Windows graphics system, an OSF/Motif

graphical user interface, and network TCP/IP compliance. This permits the system the widest range of portability to the largest number of platforms.

SATORI Applications

Prior to the development of SATORI, it was not possible to replay graphically the movement of aircraft targets and their associated data blocks across a given sector of En Route airspace synchronized to the associated voice data. This capability has the potential of improving the safety of the NAS. A discussion follows of the potential benefits from using SATORI.

Incident Reviews

Quality Assurance (QA) teams will have a tool for reviewing the situational dynamics which occurred during an incident under review. Currently, QA review of an OE involves looking at the NTAP printout, a process that is limited to the display of two to five aircraft targets and associated limited data blocks presented on paper. SATORI will provide the capability for the simultaneous display of all aircraft targets and data blocks for a given sector of airspace in a video format in sync with the associated voice track. NTAP would remain the legal tool used for assessment of the loss of separation; however, this system would facilitate investigation in determining the situational dynamics that took place prior to and during an error. It would also provide an opportunity to demonstrate "good" techniques of control and outstanding flight assists.

DYSIM Review

It will be possible to use SATORI to review performance on DYSIM problems without the use of training time on a PVD. At present there is no capability to replay a given controller's or trainee's performance on a DYSIM problem. DYSIM data are recorded on the SAR tape and would involve the same process for data reduction as that for actual ATC situation re-creation. It is hoped that, through the review DYSIM performance and OEs and the determination of the tasks that were omitted or were done incorrectly, it will be possible to identify those tasks which should be addressed in remedial training.

ATCS Performance Measures

SATORI has the capability to provide a basis for developing objective measures of controller performance rather than over-the-shoulder assessments derived from simulated scenarios. A number of measures have been developed by researchers at CAMI and the FAA Technical Center for use in evaluating ATCS performance. These measures and the system

being developed for their calculation using NAS data are described in detail in a report by Rodgers and Manning (1993).

If the performance of several controllers is recorded for a DYSIM problem, it should be possible to compare an individual's performance with group performance on the same problem. It would be possible to review over-the-shoulder assessment techniques for accuracy. This would utilize the objective measures of performance currently under development as mentioned above. These measures would provide a standard for training ATC instructors or evaluators in over-the-shoulder rating techniques and a means to objectively evaluate their performance.

OE Research Tool

SATORI will provide a research tool for investigating the tasks and controller actions involved in the commission of En Route OEs. Prior to the development of SATORI, there were no means by which the dynamics of the control situation could be understood since they were unavailable for review. With the use of SATORI, the job tasks taxonomy (Rodgers and Drechsler, 1993), the reference guide to tasks associated with the causal factor categories of the current operational error reporting system (Rodgers, 1993), taskload and performance measures (Rodgers and Manning, 1993), and the assistance of the controller involved in the error, it should be possible to identify which actions were omitted or which were done incorrectly to precipitate the error.

Design Appraisal

The FAA has a requirement to evaluate its current systems against proposed future ATC systems designs. Without a means to objectively assess the current ATC system ATCS taskload and its effect on ATCS performance, it is not possible to make a meaningful evaluation of the potential impacts of any design changes proposed to be made. With the likely increase in automation of future ATC systems, it becomes imperative that the Agency have the capability to make comparisons between proposed system designs and the current system. Without such a tool the design process will be speculative and not cost effective.

Traffic Management

SATORI will allow for the review of the impact of airspace design on the flow of traffic through a facility's airspace. Currently, SATORI is simply an animation tool which re-creates the traffic flow through a sectors airspace. The development of a simulation routine to evaluate redefined airspace characteristics is currently under development. This software would allow for the modification of airspace, airways, navigation aids, etc., to review their effect on traffic flow. Calculation of measures of sector characteristics is also under development.

Future Developments

An automatic OE detection system was implemented in domestic ARTCCs in 1984. Under this system, an alarm is triggered whenever minimum separation standards are violated by radar-tracked aircraft. Although this system allowed for violations of separation standards to be more closely monitored, it did little to assist in determining why or how such violations occurred. It is planned that SATORI will provide a means for determining the particular system component that failed. Tasks that consistently cause problems for controllers and which would be candidates deserving consideration for automation or the development of automated aids should be identified. Instead of attempting to automate all tasks, it would be better to automate only those that have the potential to give controllers problems. Research has demonstrated that humans are not well suited for passive monitoring of system performance. Performance typically improves when they are actively involved in the control of the situation. The optimal level of automation in air traffic control is yet to be determined, and the potential for over-automation should not be discounted (Endsley, 1992). The effect of automation on ATCS situational awareness deserves greater attention.

Additionally, a development is underway to attempt to utilize Continuous Data Recording (CDR) data from the TRACON environment to provide a SATORI tool for those facilities. Similar data to those available from En Route facilities are recorded at TRACON facilities and should allow for the development of a re-creation tool much like the one discussed in this report.

The development of software to allow for the assessment of performance and taskloading is currently underway. Most of the required algorithms for the various measures have been derived and are currently being converted to source code. This development will be discussed in detail in a report in preparation (Rodgers and Manning, 1993).

Once SATORI is developed and evaluated, it will be possible to accomplish the goals of evaluating system designs, over-the-shoulder appraisals, training outcomes, and measuring controller performance. Not only will the capabilities and features of SATORI provide those interested in air traffic with a valuable tool for assessing the dynamics of the air traffic situation, but additionally, and more importantly, the Agency will be in a better position to bring about effective change in future ATC Systems.

References

Endsley, M. R., (1992). *Situation Awareness: A fundamental factor underlying the successful implementation of AI in the air traffic control system.* NASA/FAA Workshop on Artificial Intelligence and Human Factors in Air Traffic Control and Aviation Maintenance. Daytona Beach, FL., June.

Federal Aviation Administration. (1991). *Air Traffic Control (FAA Order 7110.65G).* Washington, D.C.: U.S. Department of Transportation.

Rodgers, M. D. & Manning, C. A. (1993). *Performance and taskload measurement of En Route air traffic control specialist.* FAA OAM 93-in preparation. Washington, D.C., Federal Aviation Administration, Office of Aviation Medicine.

Rodgers, M. D. (1993). *Relationship between En Route air traffic control specialist job tasks and causal factor categories of the current operational error reporting system: A Reference Guide*. FAA OAM 93-in preparation. Washington, D.C., Federal Aviation Administration, Office of Aviation Medicine.

Rodgers, M. D. & Drechsler, G. K., (1993). *Conversion of the CTA Incorporated job tasks database into a formal sentence outline*. FAA OAM 93-01. Washington, D.C., Federal Aviation Administration, Office of Aviation Medicine.

The Impact of Associate Systems Technology on an Air Traffic Controller's Situational Awareness

Belinda D. Hoshstrasser & Ronald L. Small

Search Technology, Inc.

Introduction

Associate systems technology has been applied in a variety of complex domains and could prove useful in air traffic control (ATC) (Greenberg and Small, 1993). Associate systems offer many potential benefits to air traffic controllers as they build and maintain awareness of their increasingly complex airspace situation. Associate systems can provide intelligent information management based on the controller's activities and workload, assistance through adaptive aiding, automated monitoring for errors, planning, problem solving, and situation assessment. However, these functions are not without additional human factors problems. For example, how does the associate system effectively communicate high level assessments or recommended plans of action to the human controller? Will the air traffic controllers accept the recommendations of the associate system? Will controllers be distracted by a high-tech system, potentially losing situational awareness during critical moments? This paper identifies some of the potential human factors problems introduced by intelligent systems and recommends some principles that can be applied to reduce these problems.

How Can Associate Systems Technology Help In ATC?

Air traffic controllers have an ever increasing information overload. Every day more and more flights fill the skies, leading to increased air traffic density. More sophisticated equipment is being introduced into both the airliner cockpit and the air traffic control facilities (control towers, en route centers, and approach and departure control rooms) producing more information-laden outputs.

Associate systems can help in several areas. First, by managing the display of information, the controller sees the information that is needed, when it is needed, and in a format that is easily interpreted. An associate system can combine bits of information from independent sources into a more useful presentation, relieving the controller from the burden of mentally assimilating independent pieces of information to form a cohesive assessment of

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

the current situation. The controller and associate system can work together to accomplish tasks and resolve problems. Planning functions inside an associate system provide recommended plans of action. In addition, error monitoring capabilities detect and provide remediation for errors of commission or omission which are either procedural errors or erroneous intentions. Adaptive aiding functions assist through automated task allocation and execution (Rouse et al., 1987). All of these associate system functions serve to focus the controller's attention on the important tasks and information, thereby enhancing his or her situational awareness and ability to accomplish an increasingly complex set of responsibilities.

To be truly effective, associate systems must strike the proper balance between the human's responsibilities and the automation's functions. Therefore associate systems must address all the traditional human factors concerns, plus a new set of human factors concerns. Three of these concerns will be discussed in this paper:

- user acceptance of the associate system;
- effective communication between the controller and associate system; and,
- cooperation and coordination between the controller and associate system.

Stated another way, the associate must enhance the controller's situational awareness, not cause distractions. In the following sections, each of the above concerns will be identified as a set of questions. Then, principles providing possible solutions to these concerns will be presented.

User Acceptance

In general, a successful associate system addresses the problems listed below in achieving user acceptance. The answers to the following questions are impacted and determined during system design and early development via knowledge engineering, prototyping and system testing.

Problems

- Will the controller accept the advice given by the system?
- Will the controller trust the system?
- Will the controller enjoy working with the system?
- Will the associate system make the controller's job better?
- Will the controller still feel useful and necessary in the air traffic control system?

To make an associate system for air traffic controllers worth the effort of building it, the system must be accepted by the controllers. Many of the traits that are desirable in a human coworker or team member are also desirable in the behavior of the associate system. Two major factors influence user acceptance: trust and system utility.

Principles of Trust. A level of trust must be established between the controller and the associate system so that the controller has faith in assessments and recommendations provided by the system. Trust can be established through training and, more importantly, by involving the air traffic controller community in the design and development of the associate system and its knowledge bases. Building an operator's trust of an associate system is analogous to building trust between human coworkers, for example flight crew members or teammates (Taylor, 1988; Lizza et al., 1990), so the problems are not insurmountable. Also relative to promoting a level of trust, the associate system must behave in a predictable manner. An unpredictable, erratic system will not be trustworthy. System utility influences level of trust as well, and is discussed next.

Principles of Utility. If the associate system provides little or no added utility compared to the current air traffic control system, it will not and should not be accepted by the user community. Many factors influence the perceived usefulness of an associate system. There are several principles which can be employed to ensure system utility and user acceptance:

Provide valuable information. The associate system should display information that is important and necessary given the controller's current activities. It should enhance the controller's situational awareness, not detract from it.

Provide appropriate assistance. Help should be provided by the system only when it is needed or requested by the controller. Otherwise, the human controller will feel like the system is watching over his or her shoulder and taking control away -- similar to an overly involved supervisor who does not trust the actions and judgment of a subordinate.

Avoid being a nuisance. This is related to the previous two principles. Providing irrelevant information, emphasizing information the controller already knows or excessively assisting the controller, is bothersome, might detract from the controller's situational awareness, and is potentially dangerous. Imagine working with a person who is always quoting irrelevant or obvious facts and constantly reaching over your shoulder to "help" you. No one wants to work with a person who behaves this way; an associate system exhibiting such behavior is equally undesirable.

Minimize false alarms. (Don't cry wolf!) This is potentially worse than missing a minor alarm occasionally. By constantly alerting the controller of the possibility of impending danger, the controller will become desensitized to the warnings and begin ignoring them. The associate system should not over-react to situations with attention-demanding display functions (e.g., aural or visual warnings). Excessive over-reactions or interruptions will result in the controller turning the associate system off.

Don't miss alarms when collisions are impending. The associate system must detect imminent disasters without fail, providing the controller both the time and opportunity to react to the situation.

In addition to these associate system behavioral principles, there is an important design principle for system designers to keep in mind. The associate system must be perceived as an *assistant*, not a supplanter of the controller. Associate systems are intended to provide assistance to humans, not to replace humans or to reduce the job of the human to a system monitoring task.

Trust and utility are crucial for system acceptance. The style of communication between the controller and associate will help foster the sense of trust and the perceived system utility. Methods for achieving effective communication are described next.

Effective Communication Between Controller & Associate

Effective communication between the controller and associate system is of extreme importance if each member of the ATC team (human and associate) is to properly accomplish tasks. To enable effective communications, one must consider the following issues and problems.

Problems

How will the associate system know what the controller's intentions and objectives are? These intentions and objectives must be understood if the system is to provide valuable assistance.

How can the associate system best communicate information to the controller? High level assessments and plans are new kinds of information to be communicated. What is the best way to do this?

Principles for Effective Communications

Principles for effective communications can be broken down into two groups: controller to associate system and associate system to controller.

Controller to Associate System Communication Principles. First, the controller should not be required to explicitly inform the associate system of his or her intentions. If the controller was required to explicitly update the associate system's internal model of controller intentions, the system would be of little use. Instead of providing assistance, the associate system would actually create more work for the controller and thus become a burden. An associate system should observe the human's interactions with the ATC system and infer intent, just as effective human teammates do. By watching the actions of another person in context, people are often able to infer the intentions and purposes behind another person's

activities. Intent inferencing mechanisms have been implemented based on these principles (Hoshstrasser and Geddes, 1989; Hoshstrasser and Skidmore, 1991).

The second important principle for effective controller to associate system communications is the ability of the controller to either explicitly or implicitly accept or reject any plan of action suggested by the associate system. The controller has the freedom to evaluate the associate system's advice and decide whether or not to take it. After all, in today's air traffic control system, the controller is responsible for the outcome. This is likely to be true in the future as well. So, the human must have complete authority to accept or reject any associate system recommendation or assessment.

The most important point to remember in considering controller to associate system communications is that *the human is ultimately in control of the system*. It is up to the human to do whatever he or she thinks is best in a given circumstance. The associate system must then infer controller desires and intentions without placing the burden of communication on the controller.

Associate System to Controller Communication Principles. All communications with the human controller must be for a purpose, as mentioned in the section on user acceptance. Information should be communicated to the controller for the purpose of enhancing situational awareness by:

- notifying the controller of significant events;
- notifying the controller of potential problems;
- supporting the controller's tasks; and,
- recommending plans of action derived by the associate system.

Notifying the controller of significant events and problems implies that the associate system understands which events and problems are important to the controller. This knowledge must be obtained from the air traffic control community and encoded into the associate system's knowledge bases. To support the controller's tasks with appropriate information, attention must be paid to the controller's current focus of attention. When the human is focused on a particular problem or task, the associate should provide more detail relative to that particular task and eliminate irrelevant or extraneous information. Recommended plans of action must be in harmony with what the controller is already trying to accomplish, otherwise confusion can result, and situational awareness may be decreased.

Communications should be designed with one overriding concern in mind: the purpose of the associate system is to *assist* the controller. If this human-centered design approach (Rouse, 1991) is strictly applied, communications between the controller and associate system will go smoothly. Besides communication style, the associate system must behave in a coordinated and cooperative fashion with the human controller, just as any effective human teammate. The issues of cooperation and coordination are discussed in the following section.

Cooperation Between Controller & Associate

Three aspects of controller and associate system cooperation and coordination will be discussed:

- adaptive aiding through task allocation and task sharing;
- assessment and solution of problems; and,
- identification and resolution of controller errors or mistakes.

Task Allocation

Task allocation is the process of determining which part of the overall system (human, associate or both) should perform the task under consideration. There are instances when only the human should perform the steps in a task; there are times when it is more appropriate for the associate to perform the task steps; and, there are times when the controller and associate should share task steps. Determining the proper allocation for a given instance is a non-trivial knowledge engineering issue (Anoskey and Andes, 1992). But, for this paper's purposes, the human factors issues are the main concern. These issues can be addressed by answering the following questions:

- How will the controller know when a task has been allocated to the associate system?
- How will the associate system know when the controller needs assistance?

The controller needs an accurate picture of what is going on at all times. This is an important consideration in effective workload reduction through task allocation. Answering the questions above will contribute to effectively and appropriately reducing controller workload leading to enhanced situational awareness. Principles to follow when answering the above questions during the design of an associate appear next.

Principles of Task Allocation

Several principles can be applied to designing an effective task allocation scheme. First, assistance should be offered only when it is actually needed. This allows the human to remain a useful part of the system and not become a monitor of the associate system's activities.

Task allocation should be predictable and unambiguous. To help with this, task allocation strategies can be designed such that tasks may be allocated on an always or never basis, or based on the operator's workload or criticality of the situation. Each individual controller should be allowed to define his or her own preferences for task allocation. For example, a particular controller may prefer that routine or mundane tasks always be handled by the automation. On the other hand, another controller may decide to allocate routine tasks to

automation only when workload is high. Some tasks, however, will always require human judgment and should never be allocated to the automation.

Regardless of the task allocation preferences defined by the controller, he or she should be able to assume any task, even after the automation has begun the task. In order for the human to maintain a good mental picture of the situation, remain in control, and exercise authority over tasks (for which he or she is ultimately responsible), the controller may take control of any task at any time, finishing it as the associate had planned, altering the task, or even replacing it with another task entirely.

To maintain awareness of the current situation, the controller must know the status of automation-allocated tasks. This means that the controller knows when a task has been allocated to automation, when the task execution actually begins, the progress of the task during execution, and when the task is finished. It is likely that other controller activities will be depend on the status of allocated tasks.

Detection and Solution of Problems

The types of problems that an associate focuses on are problems relevant to the human operator, based on the operator's intentions. That is, the associate system follows the lead of the human controller, sharing the same concerns. The reason for this focus is that if the associate presents a solution for a situation that the controller does not recognize to be a problem, the presentation of that solution may confuse the controller. Of course, there are instances when the associate needs to inform the controller of a problem that he or she may not be thinking about, such as controller errors or sudden events; but, these problems are handled differently (as described in the section on Identification and Resolution of Errors, below). Another reason for focusing on the controller's intentions is to ensure that the associate system is supporting the controller's activities.

Associate system designers should answer the following questions and address the following issues using the ensuing principles:

- How will the associate system identify problems that are of concern to the controller?
- Will the associate system plan solutions to problems that the controller will accept? (This relates to trust and user acceptance, again.)
- How will problems and recommended solutions be communicated to the controller?

Principles of Associate Problem Solving

Assessments of impending problems made by the associate system must be meaningful and important tc the controller. This point has been mentioned before; an example will help emphasize its significance. Many individual pieces of data are available to the associate system. There are a multitude of potential assessments that can be made from the data. For instance, the associate system could tell the controller the average altitude of every airplane in the sky. However, this information would be of little interest to the controller. Consequently, the associate's assessment and planning functions must operate in support of the controller's current activities and intentions. Additionally, assessments or plans for dealing with airplanes

that are still hours away are no help to the controller when there are more immediate problems at hand. Using an internal model of the controller's intentions and the associate system's own knowledge of the current situation enables the associate system to focus its assessment and planning functions on supporting the controller.

A very important lesson learned in designing associate systems for other domains is that the human must be able to interact with planning functions (Hoshstrasser and Skidmore, 1991). It is rarely the case that a human will be in complete agreement on plans of action recommended by either an associate system or another human. People have their own perspectives on how problems should be solved based on training, experiences, and individual differences. Because of this, the controller needs the ability to modify, accept, reject, or replace a plan proposed by the associate system.

Identification and Resolution of Errors

Human errors are an unfortunate side-effect of the human's abilities to innovate and use ingenuity in problem solving. An associate does not try to prevent controller errors; rather, it identifies observed errors and resolves them so as to mitigate the serious consequences of the errors. No complex system (e.g., air traffic control) can be designed to eliminate errors. But, error consequences can and should be identified and resolved (Greenberg, 1992; Greenberg and Small, 1993).

In identifying and resolving errors without unduly detracting from the controller's current tasks, the important human factors issues of communications and user acceptance arise again. These issues are posed as the following questions and are answered by another set of principles:

- How will the system notify the controller of errors and the proposed resolution of those errors?

- Will the controller accept the system's assessment and proposed resolution of errors?

- Identification of errors presents a unique problem for the associate system: How will the system distinguish innovative activities from errors?

Principles of Error Identification and Resolution

The issues of error notification and controller acceptance of suggested remediations can be solved in the same way as the detection and solution of problems described earlier. Errors should be pointed out to the controller only when the error has important consequences. For example, if an aircraft has been assigned to an altitude of 14,000 feet, but is actually at 14,100 feet, this is not an error. The consequences of an aircraft being 100 feet off of the assigned altitude are negligible when 300 feet of variance is typically allowed.

The controller should be able to communicate to the system when he or she is engaged in innovative or unusual behavior that the associate may perceive as errors. Situations inevitably arise that system designers did not anticipate. The controller needs the flexibility to deal with these situations as he or she sees fit -- without the associate system insisting that

the controller is committing an error. When such situations arise, the controller needs to be able to tell the associate system, "It's okay -- I know what I'm doing."

All of the principles espoused so far can be combined into an overall set of principles for interacting with associate systems so that these associates will enhance the human's situational awareness and job performance. These principles can and should guide the design of these systems, and are described next.

Principles Of Interaction For Associate Systems

The problems identified in the preceding sections lead to a higher level set of principles that are applicable to all associate systems. These principles derive from the more general human-centered design guidelines described in the next section.

- The human has ultimate control and can override the associate system at any time.
- The associate system must follow the lead of the human operator. The associate monitors and supports the human, not the other way around.
- The behavior of the associate system must be predictable.
- The operator should have the option of turning off part or all of the associate system. For example, error monitoring may be turned off when the operator is involved in innovative, unconventional behavior. The associate system should still observe the activities of the controller and the events taking place in the outside world, maintaining its internal states and models, and presenting information useful to the controller.

Human-Centered Design

Along with a set of principles of interaction, a human-centered design approach should always be taken with associate systems. A human-centered design philosophy will ensure that the associate system meets its design goals of supporting the operator. After all, the main purpose of the system is to support the human operator in his or her tasks.

The inherent abilities of humans and computers should also be primary concerns in designing an effective associate system. Designers must consider what kinds of tasks humans do best and worst, and what tasks computers do best and worst, (Fitts, 1951). For example, computers are tireless, ever-vigilant monitors of data. Humans, on the other hand, get bored, make mistakes, and are generally unhappy when their only responsibility is monitoring for a change in data, especially in cases where changes are slight and infrequent. When changes occur rapidly and unpredictably, humans can become overloaded.

Principles of social psychology should also be taken into account when designing an associate system. The system should be designed to exhibit the same characteristics as a good coworker or team member.

Conclusions

Associate systems technology offers may potential benefits in complex human-machine systems, such as air traffic control. However, to make the most of the advantages provided by associate systems, the human user must be considered throughout the design of the system. This can be ensured by using a human-centered approach to design and development, and by establishing a set of Principles of Interaction between the human operator and associate system. The use of these design principles will result in systems that enhance the human's situational awareness, job performance and job satisfaction, all of which are crucial to operating increasingly complex systems in increasingly complex environments.

References

Anoskey, A.M. & Andes, R.C. (1992). *Guidelines for Adaptive Aid Design: A Review of the Literature.* STI-TR-9103-001. Norcross, GA: Search Technology, Inc.

Fitts, P.M. (1951). Engineering psychology and equipment design. In S.S. Stevens (Ed.), *Handbook of Experimental Psychology.* New York: Wiley.

Greenberg, A.D. (1992). Monitoring for hazard in complex systems. *Workshop on Expert Judgment, Human Error, and Intelligent Systems at the 10th European Conference on Artificial Intelligence*, 169-187. Vienna, Austria.

Greenberg, A.D. & Small, R.L. (1993). Improving human reliability through error monitoring. *Proceedings of the IEEE Symposium on Reliability and Maintainability.* Atlanta, GA.

Hoshstrasser, B.D. & Skidmore, M.D. (1991). The design of an object oriented intent inferencer for integrated systems. *Presented at the 1991 Digital Avionics Systems Conference.* Los Angeles, CA.

Hoshstrasser, B.D. & Geddes, N.D. (1989). OPAL operator intent inferencing for intelligent operator support systems. V.L. Shalin and G.A. Boy (Eds.), *Proceedings of the IJCAI-89 Workshop in Integrated Human-Machine Intelligence in Aerospace Systems*, 53-70. Detroit, MI.

Lizza, C.S., Rouse, D.M., Small, R.L. & Zenyuh, J.P. (1990). Pilot's Associate: An evolving philosophy. *Proceedings of the Workshop on The Human-Electronic Crew: Is the Team Maturing?* Ingolstadt, Germany.

Rouse, W.B., Geddes, N.D., & Curry, R.E. (1987). An architecture for intelligent interfaces: Outline of an approach to supporting operators of complex systems. *Human-Computer Interaction*, 3(2), 87-122.

Rouse, W.B. (1991). *Design for Success: A Human-Centered Approach To Designing Successful Products and Systems.* John Wiley and Sons, Inc.: New York.

Taylor, R.M. (1988). Trust and awareness in human-electronic crew teamwork. *Proceedings of the Workshop on The Human-Electronic Crew: Can They Work Together?* Ingolstadt, Germany.

Weather Services

- Individual Differences in Weather Situation Awareness and Assessment
- Situation Awareness in Marginal Weather Conditions: Do Graphical Presentations Help?

Individual Differences in Weather Situation Awareness and Assessment

C. Elaine McCoy[1], JoAnne Kollross Woleben[1] & Philip J. Smith[2]

[1]University of Nebraska at Omaha
[2]The Ohio State University

Introduction

In our highly complex and sophisticated national aviation system, discrete information is provided to members of task groups with differing operational goals and legal responsibilities. While there is a common ground, much information including weather data and briefings remains more or less available to specific subgroups. Air Traffic Control, flight crew, and dispatch are three groups performing unique tasks that require cooperative planning and cooperative execution of operations to insure safety in the aviation system. Each is a member of a technologically-dependent task group that has both common and competing goals.

Situational awareness is still being defined in its application to civil aviation. Considered as an individual performance issue, situational awareness training and operational issues can be easily viewed as applicable to individual members of groups such as pilots and air traffic controllers. Successful flights, however, require that individuals and groups of individuals must selectively and cooperatively communicate relevant aspects of their unique awareness of the situation-at-hand. In order to perform at an optimum level, the pilot must be aware of the parameters which constrain the other participants of the larger system while simultaneously delineating his or her own intended path.

Fundamentals

Two general, fundamental characteristics of situation awareness are:

- situation awareness is always dynamic; and
- situation awareness requires a human player who is oriented toward a task.

Situation assessment goes beyond the perceptual act. Situation weather awareness in civil aviation incorporates cognitive aspects of context assessment that integrates knowledge, experience, and presence.

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

Cognitive activity is required to build models and, in complex situations, cognitive simplification strategies may be used to aid in comprehension and to evaluate the alternatives in that context. To distinguish situation assessment from situation awareness: situation assessment requires cognitive model building activity and situation assessment may utilize cognitive simplification strategies to build such models and evaluate alternatives.

Collective Utilization. The term collective utilization addresses the limits on the ability of the individuals of the technologically-dependent task group to share their unique experience and awareness yet provides for a common ground of communicated or common understanding. Collective utilization of experience (awareness/assessment) occurs in a dynamic context. To truly delineate situation awareness at a fundamental level, issues of human motility and spatial and temporal relations must be addressed.

Spatiality of Situation. There is a difference in "spatiality of situation" and "spatiality of position." The more fundamental spatiality is that of situation which may seem contrary to our everyday view. This is not a "determinate position in relation to other positions or to external coordinates, but the laying down of the first coordinates the anchoring of the active body in an object, the situation of the body in face of its tasks" (Merleau-Ponty, 100). This "spatiality of situation" is the engaged human performer who both defines and is defined by the tasks in which he or she engages. The human performer whether pilot, controller, or dispatcher has a delineated and unique responsibility. This is a fundamental orientation which is prior to an identification of, for example, navigation/geographic orientation.
In our application, the controller has a unique responsibility and orientation toward separation of aircraft and flow control; the pilot is toward the operation of a specific aircraft – movement, tasks, destination; dispatch is toward air carrier operations as a whole. The consciousness of each is polarized towards its aims and these projects polarize the world and bring to view signs which guide the person's actions. The human performer is, thus, the always tacitly understood third term for all figure-ground relationships and each figure is present against the horizon of external and bodily space. (Merleau-Ponty).

Significance requires a point of view, a certain distance, and a certain direction (Merleau-Ponty, 429). The dimensional movement of aircraft across a scope is significant from the controller's stance. For the pilot, an instrument approach – the approach path, the chart, the instruments, the time – is significant in relationship to the airport and runway. However, to the extent that the flight of the aircraft is designated by the pilot, the perceptual field of the pilot subsumes the aircraft incorporating it as part of the bodily stance from which the approach is made. The pilot flies (McCoy, 1985). Such a perceptual field is the fundamental "place" or moment of collusion with the world (Merleau-Ponty) that founds a fundamental aspect of situation awareness.

Spatial and Temporal Metaphors. The movement toward something in space is also a movement toward a temporal horizon. The movement is not just from one place to another but is also from now to then. While there is no direct perception of space or time, specific systems utilize specific metaphors which vary in a number of aspects (Mickunas, 1977).

The aviation professional's situation assessment will be made presupposing a particular spatial/temporal relationship. Both pilots and controllers operate in domains in which spatial/temporal emphasis-shifts occur for individuals as the context of performance changes. This difference in emphasis also contributes to the differences that can be seen in system participants' assessments of how the weather impacts a flight or flights.

The Function of Time. Space is temporalized for much of the world of aviation. That is, time determines the spatial situation. Flight parameters are determined by time. I have x hours or minutes of fuel remaining and that determines my options – the space of my choices to land, refuel, to stay aloft. Time places parameters on the spatial environment; it selects or chooses the range of possible spatial alternatives. In effect, time expands or contracts the spatial situation. Flight, the domain that collapses time, clearly maintains the priority of time in flight replanning. In a pilot's speech this cogently understood phenomenon is apparent: "We will arrive in Orlando in 3 hours" not in 1200 miles. ASRS Report #75202{1} illustrates that time is treated as primary for a pilot in this phase of flight (emphasis added):

> Unforecast weather combined with extensive delays, extensive vectoring ...information we got regarding times involved was unrealistic. Approaching destination we were told to deviate east from the arrival due to a line of thunderstorms...we were held...then another vector to pick up a radial into MIA then back up to FLC. But when we arrived on SBY we were again put in a hold with EFC 10 min. At this point we told them we could not take much more added delay. This is when we were told 12 aircraft ahead of us and they wanted 20nm in trail from holding fix. This obviously meant 35 to 40 minutes added so the captain elected to land at Nassau (not a regular alternate) and refuel....In this case even though we tried for the last hour or so before scheduled arrival to judge how the situation was progressing we were told that "just a little more."

We could have been in bad straits had Nassau also closed. Nassau became an unscheduled stop that was within range--distance defined by the rate/hours of fuel remaining. The pilot responses primarily refer to time. A hold for a pilot is to be essentially fixed about a point in space while time passes. The indicated directions from ATC, however, focus on space such as the "distance in trail from a holding fix" in the above report. The pilot implies that inadequate information was provided by ATC concerning the length of delay which also kept the flight crew from assessing the situation and gauging the risks earlier.

Pilots may shift space to the foreground when, for example, the context of the thunderstorm requires local maneuvers in the vicinity of an airport. Then the consideration becomes one of position relative to or oriented to the thunderstorm and runway. ASRS report #147629 illustrates the effect of a thunderstorm on approach when space assumes the foreground for both ATC and the pilot although with differing significance:

> There were some thunderstorms on the arrival route. The first controller was most helpful in letting us deviate... The next guy wanted to know what we were doing. He gave a turn to return to the published arrival route that would have put us right in a thunderstorm. ...It took two requests to get him to let us deviate. Turning final we were below the base of the clouds but could not see the field because of a heavy rain shaft just outside the marker. The cell producing the shaft had heavy contour on the radar. I was making some strong assumptions, contour ... high bases, small well-derived rain shaft. Are not those all signs of a pending or possible microburst? We could see that if we would slide just left of the rain shaft, and maybe be south of the outer marker by only a half mile we could avoid the rain shaft, stay VFR and land. I had the first

officer tell the tower our intentions. A different voice came on the tower and broke us out of the pattern...

The perspectives and information available to the flight crew differed greatly from that of ATC. There was little or no recognition by flight crew or ATC of being members of a task group with common goals. The flight crewmember expresses no thought of potential traffic conflict while ATC seems determined to have the aircraft conform.

Although ATC maintains aircraft separation in both time and space, it may be argued that for ATC space is used to contain aircraft. Aircraft are held in a sector or at a fix or permitted to approach or depart an airport. The orientation of ATC is toward the physical separation of aircraft. The controller considers other aircraft positions, geography, airspace, and weather. Weather intervenes and is addressed as a disruption of the use of space.

In the following example, both the pilots and ATC focused on space but again with different concerns – the pilots were aware of their imminent encounter with a thunderstorm and the risk inherent in its proximity while ATC was focused on the relative position of the aircraft to jet routes and to restricted airspace. ASRS Report #119515:

We departed San Francisco for Las Vegas....we were aware that thunderstorms existed east of the Sierras. Approaching the mountains we advised we would soon have to begin deviating. OAK advised us that J92 was "closed" due to thunderstorms and we should expect to be rerouted via Milford, UT. We were on initial vector toward Milford as we crossed J92 S of Coaldale. ZOA advised J92 was now open and gave a vector heading toward Beatty. We replied we would have to hold our current heading to go around the East side of a large thunderstorm on J92. Shortly, center (LAX by now) asked how much further East we had to go. We replied 15 miles. LAX said that in 7 miles we would reach boundary of a restricted area. In the time it took to say all this, LAX said 'turn now' to a heading that would put us right in the thunderstorm, according to LAX no other heading would keep us out of the restricted area. Captain declared an emergency and flew headings to stay out of thunderstorms, but apparently we entered restricted areas....

"In the time it took to say all this" reaffirms the conjoined movement toward temporal and spatial horizons. This spatial emphasis shifts to the temporal for the controllers when rates of closure have priority or when controllers run out of space in which to vector aircraft.

Pilots and controllers, although required to work and problem-solve as task group members, frequently begin with contradictory perspectives so fundamental as different foreground/background relationships of the spatial/temporal metaphors that are in use. Not only do different members of the task groups have different information, but their focus or framing of the situation can be counterpoised. Furthermore, "space" can become the space of the display world of the scope rather than the pilot's space of movement. ATC computer models of the world can lead the human performers approaching the task through that restricted focus to lose the larger picture for the other task group members.

It should not be surprising that little recognition exists among ATC, pilots and operations (dispatch) of a unified task group. They have been trained to exhibit differences in thinking and to employ as tools displays that emphasize spatial-temporal differences.

Differences in Situation Assessment

NASA-Ames/FAA Study. In two empirical studies of flight planning as performed by 30 airline pilots and 27 dispatchers, Smith, McCoy, and Layton found wide variations in the weather models developed by different individuals (even though they all had access to the same data). As an example, in response to one weather scenario, the following evaluations were voiced:

> "I don't like the easterly route because if I get off and the line fills, I really don't have a good alternate choice other than going back up to Amarillo." "The eastern route is a pretty good route. I don't have any problem with that." "My thoughts remind me of the Southern crash and I know just a little bit about that. That area is very susceptible, doesn't give you much space in that particular area to go through the red...and I have no desire to go through that and knowing that, could fill in no time at all. That almost looks like that could be a front although its not indicated. You're going to develop hail probably in that area and so forth which specifically could give you severe danger in that area." "There are four areas of severe thunderstorm activity. It's what I'd still call scattered to widely scattered."

Such differences in the mental models (Gentner and Stevens, 1983) of the weather, and in evaluations of alternative routes, were common in these two studies. A number of factors appeared to influence such assessments.

Selective attention to data. Not all individuals looked at the same data displays. Furthermore, even when looking at the same display, different people focused on different parts of the display.

Differences in Training and Experience

Different problem-solving strategies (Fraser, Smith and Smith, 1990; Hayes-Roth and Hayes-Roth, 1979; Hoc, 1988; Suchman, 1987; Wilensky, 1983.) Different people used different approaches to generate and evaluate alternative flight plans. These approaches influenced the selection of data for viewing and the interpretation of data.

Susceptibility to the computer's portrayal of the situation. The computer allowed the individual to watch the aircraft move long routes as forecast weather was shown. In response to such displays, some individuals failed to think about the uncertainty associated with the forecasts. In short, they got deluded into accepting the computer's simplified model of the world. It appears that as more sophisticated displays are made available, people are more likely to appropriate or to become aware of the wrong world, the computer's world. The computer creates not only a tool that replicates components of situation awareness but the computer generated information becomes accepted as the sum total of situation awareness.

Such findings highlight that situation awareness is strongly determined by the cognitive processes involved in accomplishing a goal such as developing a new flight plan. Furthermore, they indicate that the subjects' cognitive processes are strongly influenced by the design of the computer system that is providing access to information.

Volpe National Transportation Systems Center Report. Individuals within the system operate on diverse sets of weather information. Evaluating performance across even a particular set of participants, such as air carrier or 135 pilots, is difficult since the information available to each varies greatly from company to company and even from trip to trip. This variation complicates the ability of the other system participants such as ATC to develop an understanding of the reliability and validity of the weather models with which the air carrier pilots are operating. Within one sector or approach area several different aircraft crews can be operating using company-provided weather information that is different from that of surrounding aircraft and ATC.

A recent VNTSC report by Turner and Huntley (1991) of 17 major and regional carriers about weather information dissemination practices to crews confirms this variability. The regulation regarding this requirement is interpreted differently: some failed to provide NOTAMS and SIGMETS; more than 50% provided additional information on stations listed in the flight plan; and only one provided SAs for stations along the route of flight. The presence or lack of SAs could significantly affect the development of weather models with regard to trend information or the available airports along the route should an emergency develop.

Recurrent weather training has not been specified by the FAA. Crew training – while specified for initial, transition and upgrade training – is not specific regarding recurrent training. (Turner & Huntley, viii) Without recurrent training, idiosyncracies develop in utilization of weather information thus promoting further divergence of weather models at the individual level. Furthermore, performance in this area can deteriorate radically. Twenty percent of air carrier pilots in the previously discussed NASA/FAA study (Layton, Smith, McCoy) could not read a traditional wind chart--direction or velocity.

Differences and Perceived Risk Management. A recognition of increased risk should be considered as evidence of a degree of situation assessment. Lofaro and Smith suggest, "Operational risk is the crew's perception of risk probability and outcome aspects of current situation with a focus on negative outcomes." As can be seen in the previously cited ASRS narratives, awareness of risk is, at the operational level, an integral component of situation awareness.

Frequently, a pilot's assessment of risk depends upon ATC or dispatch providing information that accurately portrays the situation in a timely fashion. The lack of an accurate portrayal of the situation or the distortion of the potential time frame (delays) prevents the pilot from accurately assessing the risk of continuing an existing plan.

ASRS Reports

In #175627:
The flight was routine until inbound to Keating for Milton 6 Arrival. At that point ATC gave us clearance to hold at MIGET . Fuel was not a problem at this time although ATIS showed LGA weather was worse than forecast but well above

minimums. After 1 1/2 turns in holding ATC cleared us to hold 270 degrees heading for spacing. After holding 270 heading for about 50 miles ATC informed us that weather on MIP-6 Arrival was bad (Thunderstorms) and cleared us to north to Rockdale for an arrival from the northwest. This route would add 150 miles ...enroute to Rockdale we were given more vectors off airways which continued to eat into fuel remaining. ATC also dropped us to lower altitudes and slowed us which added to the problem. Several times I queried ATC about direct routing or further delays and was assured they would turn us towards the airport shortly. At Bridgeport VOR, fuel was 4500 lbs and I declared minimum fuel and requested no further delays. ATC handling after that point was normal....LGA weather was 6 overcast ,1 1/2 rain/fog. This incident was caused by creeping delays and constant modification of plans by ATC. They kept dangling the carrot, leading us to believe normal routing to the airport was imminent, then vectoring us off route for spacing. Had I known fuel would be that low at LGA I would have landed at Albany.

Dispatch management: dispatcher: pilot. In the following case, a dispatcher (the reporter) determined risk sufficient to assign an alternate associated with his individual model of the weather for a particular area. Despite the position of a dispatch manager, the dispatcher insured that his view or model of the weather was provided the pilot component of the task group who was dependent upon the dispatcher for information that was relevant to the pilot's operational domain.

In #56829:
I advised the Flight Dispatch Manager (FDM) of my intention to add alternate MKE .... The FDM again stated that ...I did not have justification to add an alternate to my flights...I restated, "it is my belief that there will be more than slight chance of thunderstorms...I was concerned with deepening trough, lifted Index, K index of -1 over 37, vortex centered in Western Ill , moisture The dispatcher believed that pre-frontal, nocturnal convective activity would occur. After an extended argument, the dispatcher unwillingly removed the alternate.

The captain called wanting to know where is his alternate. I repeated prior conversation...The captain replied "It is my ___, add alternate." After landing ORD, the captain called and confirmed my predictions to be correct and thanked me for making him aware of my thoughts.

National Center for Atmospheric Research Icing Briefing Task Analysis. Pilots can initiate contact with Flight Service prior to and during flights to enhance their understanding and development of weather models. This service is utilized largely by general aviation, but also is called upon by commercial operators. Frequently, the enroute contact is initiated after a discrepancy is noted between anticipated conditions and actual conditions encountered.

Icing conditions are high risk conditions. In a study involving Flight Service Station personnel who brief pilots regarding icing conditions (McCoy, Biter, Sand), briefers developed carefully constructed icing models and their own trend hence forecast models. The briefers were observed at their actual workstations completing simulated briefing tasks for both Preflight and Flight Watch positions. Individuals accessed from three to fifteen different information sources to form their models for both types of briefings. Some differences also occurred in the priority a briefer gave to a particular source depending on the whether the task was Preflight or EFAS. The Flight Service Station personnel uniformly exhibited a high

degree of identification with the pilots of the task group. The briefers indicated that they attempt to recreate this "picture" for the pilots whether it is via telephone or radio. The pilots then incorporate the briefers model into their own analysis.

In the study, responses to questions regarding how presently available information might be presented more clearly or made easier to use, briefers responded in ways that suggest their concerns about carefully forming and conveying their models of the weather:

> "To me, a picture is worth a thousand words. I like to look at a map. I like to look at a picture then once I have it in my mind then I'll go to the product in front of me – the current SA's, the current winds aloft, the pilot reports – then, I'll give that to the pilot."

> "If the computer could give you a picture of the pilot's flight path and the clouds and the icing... with the pilot going form A to B, and the bases of the clouds in relation to the freezing level, and then I could see where the icing is...."

> "Once you've got the picture of what is actually happening reference to all the advisories – we get a lot of advisories that cry wolf – so once you get an idea of where the pilot reports are coming from and the activity during the shift, you tend to tunnel into the areas and have a little more confidence...it evolves during the shift."

The briefers utilize information from the other pilots in the system via PIREPS as a major component in formulating their models of the weather which they in turn provide to the pilot being briefed.

VFR into IMC. In a 1989 Ohio State University study of continued flight into Instrument Meteorological Conditions, Rockwell and McCoy found that VFR pilots flew into deteriorating weather equally well in a computer simulation as their counterparts did in accident statistics. Individuals had difficulty in developing a model of the weather that incorporated the weather's potential impact on the flight path. The significance of the weather information as it pertained to the time of the flight, arrival time and route of flight even including airborne updates seemed lacking.

In an informal 1988 Ohio State University Survey, (McCoy) 164 pilots indicated that weather which put the aircraft at risk was the major reason for enroute replanning. One DC-9 pilot reported an incident in which a large snow storm had hit the East coast. He had been instructed to hold at Philly until the snow removal crews were off. The pilot's comment, "our fuel was running out." The flightcrew checked the alternate Baltimore and Washington both of which were down, and opted for Buffalo. Airport requirements (time for snow removal) and solutions or ATC needs (workload, separation) and solutions, even when legitimate, are not always acceptable for all task group members.

Attitudes. Task group members develop perceptions about the attitudes of the others which may or not be correct but which will impact their ability to coordinate a successful mission. Education and training could provide a better understanding of the roles of the other members of the task groups and the factors limiting their cooperation.

In #74310, the pilot suggests:

"Better weather radar in the terminal control area might give the controllers a better idea of what we are painting on our aircraft radars. Also, the controllers must understand that at least some of us pilots will not penetrate thunderstorm cells."

In #147629, the pilot perceived a high degree of risk in continued flight through the weather, but also made assumptions about the need for ATC to break off his approach, "My respect for thunderstorms is stronger than my pride/ego, and I really think the tower approach control people need to understand what thunderstorms can do to aircraft performance."

Part 135. Part 135 flights are provided less organized system support. There generally is little in terms of a sense of task group coordination or planning except by the larger freight haulers or the handful of exceptionally well-managed charter operators. The following excerpts certainly suggest that much of Part 135 operations remains an area than needs attention. Remember, these are for-hire, commercial operators who have failed to achieve an acceptable level of situation awareness.

In #163735, an obviously busy 135 pilot reports:

Called dispatch for weather... called the gate... safety briefing to paxs. checklists.. set up my nav radios.."looking out the window I saw exactly what I expected, a small town at night...After clearing the runway we realized we were at the wrong airport."

In #160993, another 135 pilot comments:

"My customers..were anxious to go...I never did find my exact location on the map."

Implications

If the goal of understanding situation awareness/assessment better is to improve safety in the national aviation system, it would seem that a major direction for future research is to identify improvements that can be made in training and to develop requirements for regulatory changes that could enhance task group performance and the accomplishment of the mission of safety of flight.

Individuals must selectively communicate relevant aspects of their unique awareness of the situation-at-hand and be aware of the parameters which constrain the other participants of the group while all accepting the mission as a common goal. Klein, Zsambok & Thordsen call a similar characteristic identified in military units "team identity." Although preliminary research suggests that dispatch-and-pilots and FSS briefers-and-pilots seem to identify the common goals of the successful flight and work well as task group members, far too often ATC-and-pilots continue to exhibit a lack of this critical safety component.

Operations Issues. All members of the national aviation system charged with assisting in the successful completion of a flight – ATC, operations, flight crew – can begin to explore the implications of understanding themselves as members of a task group with an overarching common goal rather than as individuals limited to their immediate domain.

System Design Issues. The design of the systems of communication and information dissemination, effects of displays and automated systems, and problem-solving procedures should be examined from the more global perspective of the task group.

Training Issues. The following can be studied immediately to determine how the task group approach to situation assessment and operational risk assessment might enhance system safety.

Train individuals as task group members. Train dispatch and flight crews as task groups. Increase ATC awareness of their role in cooperative problem-solving. Increase understanding of individual members of the constraints of training, equipment, workload, and responsibilities of the other members of the group. Determine a reasonable degree of cooperation and support to expect.

Train members to utilize automated systems to the benefit of the group. Identify potential "traps" of displays, e.g., taking the display or scope as the real world. How does the computer influence an individual's assessment of the context--the situation?

Identify strengths such as discrete information that should be shared with other members, and train to encourage more global exploration of solutions.

How can distributed problem-solving avoid relevant data being missed when too large a data space is made available?

Determine the extent to which spatial-temporal emphasis-differences in training and displays disrupts the safe resolution of problems at the mission level.

Determine the value of training with operational risk management as an explicit component of the training program.

References

Fraser, J.M., Smith P.J., & Smith, J. W., Jr. (1990) *A catalogue of errors* (Tech. Report) Columbus, OH: The Ohio State University Department of Industrial and Systems Engineering.

Gentner, D. & Steven, A.L. eds. (1983) *Mental models.* Hillsdale, NJ: Erlbaum.

Hayes-Roth, B., & Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive Science, 3* (4), 275-310.

Hoc, J.M. (1988). *Cognitive psychology of planning.* London: Academic Press.

Klein, G., Zsambok, C. E., Thordsen, M. L. (in press) Team decision training: Five myths and a model. *Military Review, 4,* 1993.

Layton, C., Smith, P.J., McCoy, C.E. (in press) Design of a cooperative problem-solving system for enroute flight planning: An empirical evaluation. *Human Factors*

Lofaro, R. J., & Smith, K. M. (in press) The role of LOFT in CRM integration. In R. Jensen, (Ed.) *Proceedings of the Seventh International Symposium on Aviation Psychology.* Columbus, OH: The Ohio State University.

NASA Aviation Safety Reporting System. (1992). Part 121 & 135 In-flight Weather Encounter Reports. Search Request No. 2733. Mountain View, CA.

McCoy, C.E. (1988). *Survey of flight replanning incidents.* Typescript.

McCoy, C.E. (1985) The role of fundamental expression in instrument flight performance. In *Proceedings of the Third Symposium on Aviation Psychology.* R. Jensen, Columbus, Ohio: Department of Aviation. The Ohio State University.

McCoy, C.E., Biter, C., & Sand, W. (1992) *Human factors icing briefing task analysis and interview.* Typescript. National Center for Atmospheric Research.

Merleau-Ponty, M. (1962; rpt. 1978) Trans. Smith, C. *Phenomenology of Perception.* London and Henley: Routledge & Kegan Paul.

Mickunas, A. (1977). *From space to time metaphors.* Symposium on Culture University of Chicago.

Rockwell, T. H., McCoy, C. E., & Potter, S. (1989). General aviation pilot perception of detrimental weather conditions. In *Proceedings of the Fifth Symposium on Aviation Psychology.* Ed. R. Jensen, Columbus, Ohio: Department of Aviation, The Ohio State University.

Suchman, L. A. (1987). *Plans and situated actions: The problem of human machine communication.* New York: Cambridge.

Smith, P. J., McCoy, C. E., & Layton, C. (in press) An empirical evaluation of computerized tools to aid in enroute flight planning. In *Proceedings of the Seventh International Symposium on Aviation Psychology.* Ed. R. Jensen. Columbus, Ohio: Department of Aviation, The Ohio State University.

Turner, J., & Huntley, M. S. (1991) *Sources and air carrier use of aviation weather Information.* DOT-VNTSC-FAA-91-1; DOT/FAA/FS-91/1. Washington, D.C.: Flight Standards Service.

Wilensky, R. (1983). *Planning and understanding: A computational approach to human reasoning.* Reading, MA: Addison-Wesley.

FOOTNOTES*********************************

1. The descriptions and reports provided voluntarily by system participants are explanations which are not only activities in which the reporters engage after-the-fact in order to make sense of the world, but also are explicit attempts to justify actions that are deemed to be inappropriate or incorrect within the situation. Although claims must be limited regarding ASRS reports, the narratives do provide insight into system participants' perception of the events.

The ASRS narratives in this report are selected from 104 Part 121 & Part 135 reports of in-flight weather encounter reports. Sixty-five reports reflected Part 121 operations and thirty-nine reports, Part 135 operations. In 27/51 relevant Part 121 reports and in 20/32 Part 135 reports weather was the primary cause of a perceived a need to report. Since the ASRS data base is encoded in a system of abbreviations unique to its system, references are for the most part paraphrased or summarized rather than quoted. The numbers provided are the ASRS Accession numbers.

Situation Awareness in Marginal Weather Conditions: Do Graphical Presentations Help?

Jennifer L. Dyck & Richard D. Gilson

University of Central Florida

Introduction

Although situation awareness itself is not well understood, there is general recognition of its cognitive complexity. Cognitive components or processes that have been noted as being important contributors to situation awareness include long-term memory, working memory (Endsley, 1988), individual differences (Endsley, 1988; Bolstad, 1991), attention (Endsley, 1988), experience (Garland, Stein, Blanchard, & Wise, 1992), prospective memory (Cohen, 1989), the ability to integrate knowledge from different sources (Sarter & Woods, 1991), and the influence of mental models or schemas (Sarter & Woods, 1991). Most research in situation awareness has focused on one or more of these cognitive contributors, though typically not all of them, and the research has studied fighter pilots or air traffic controllers.

Pilots have many content domain areas that they must be aware of, including communicating and coordinating actions with the crew, communications with air traffic controllers, navigational demands, performance of the aircraft, presence of enemy aircraft (for fighter pilots), and weather conditions. This research focuses on the somewhat neglected domain area of weather conditions with the view that weather is an important factor for all pilots.

The cognitive components and processes involved in situation awareness of weather are reflective of those involved in other, more studied areas of situation awareness. For example, Sarter and Woods (1991) have noted the similarity between situation awareness and mental models. We have seen this influence of schemas or mental models on situation awareness in a preliminary study we conducted. General aviation pilots were presented with scenarios containing marginal weather conditions and asked to make flight decisions. Qualitative examination of the data revealed that a number of pilots classified the weather scenario as a "typical Florida afternoon, with convective activity building into potentially severe thunderstorms in the late afternoon." In actuality, the weather scenario used was not "typical" in that the convective activity occurred much earlier than usual in the day, due to the presence of a low pressure area. Pilots who incorrectly used their mental model or schema of summer Florida convective activity made potentially dangerous flight decisions.

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

Sarter and Woods (1991) also note the importance of temporal awareness. Good temporal awareness can facilitate both diagnosis and prevention of problems. This is also related to the importance of prospective memory, or memory for future acts (Cohen, 1989). Weather is dynamic, and thus inherently temporal. It is important for pilots to assess the current weather, look back at previous weather patterns, predict future weather conditions, assess the accuracy of the forecast of future weather conditions as those weather conditions evolve, and plan appropriate actions based on current and future weather conditions.

For the general aviation pilot, two of the most common methods for obtaining weather briefings are through DUAT (Direct User Access Terminal) or an FSS (Flight Service Station) briefer. DUAT briefings are coded weather briefings that may be accessed with a modem and a personal computer (George, 1990; Schuyler, 1991). FSS briefings are obtained by telephoning an FSS briefer and obtaining a verbal report from the briefer over the phone. Both methods are textual in nature, although charts or graphs, such as radar summaries, may be obtained through DUAT for an extra fee. Given the weather conditions in a text format, like DUAT and FSS briefings, pilots must be able to translate the text into a "mental picture" of what the weather conditions are like; that is, they must be able to visualize the weather conditions they will encounter during their flight.

Pilots may be helped in visualizing weather conditions if they receive weather information in an "already visualized" format; in other words, in a graphical or chart format, rather than a text format. The structure or format of the information that is presented seems to influence the ease with which that information is processed, as suggested by researchers in situation awareness (Garland et. al., 1992; Tenney, et. al., 1992) and in cognitive psychology (Glenberg & Langston, 1992; Glenberg & McDaniel, 1992; Mandl & Levin, 1989; Mayer & Gallini, 1990). Research in cognitive psychology on the comprehension of information presented in textual and pictorial formats shows that pictures are well-suited to convey spatial information and that pictures can increase comprehension of information, but that the advantages of pictures also depend on the types of pictures used and the relationship between the pictures and text presented (Levin, 1989). Research is needed to determine if this general principle of pictorial or graphical superiority holds in the domain of aviation weather.

Hansman and his colleagues (Wanke, Chandra, Hansman & Bussolari, 1990; Wanke & Hansman, 1992) have begun work in this area, specifically looking at graphical display of windshear alerts. Pilots were asked to fly various scenarios in a simulator, during which they received microburst windshear alerts in one of several formats. When pilots received graphical alerts, they made fewer incorrect decisions than if they received verbal or text alerts. We have also examined the issue of text versus graphical presentation of weather conditions; however we have done so by using pre-flight weather information for general aviation pilots. Our research extends the work of Hansman by examining presentation format of weather in a larger time scale situation; that is, awareness of microbursts and subsequent actions must be made in a matter of minutes, sometimes seconds, while awareness of weather conditions and subsequent actions for an entire flight may be on the order of hours. This time difference may be a critical element in understanding situation awareness of differing types of weather conditions since the temporal dimension of situation awareness is important (Sarter & Woods, 1991). Our research also differs from Hansman's in that we do not present one critical hazardous weather element, but rather present an entire scenario of marginal weather conditions with no one critical hazard. This addresses a different aspect of situation awareness: that of conditions slowly evolving over time rather than the awareness of a sudden change in the situation. Recommendations for structure or format of weather presentations may differ based on the two factors of immediacy and severity of the weather conditions.

In the study presented below, general aviation pilots were presented with a weather scenario in textual or verbal formats, or each combined with a graphical format. After studying the weather scenario, the pilots were asked to make flight decisions for those weather conditions. We expect that pilots who receive the graphical weather format will make more appropriate flight decisions than those pilots who do not receive the graphical format.

Method

Subjects and Design. Forty pilots who were students at Embry-Riddle Aeronautical University (ERAU) in Daytona Beach, Florida served as subjects. The pilots were instrument rated and had a mean of 279.95 total flight hours (s = 109.99), and 42.90 flight hours of instrument rated flying time (s = 31.85). The largest percentage of their flying hours were accumulated in Florida (M = 76.87%, s = 30.25%). The pilots were also familiar with a Cessna 172 aircraft, which was the designated aircraft used in making the hypothetical flight in the weather scenario. Subjects were randomly assigned to one of four types of weather formats: 1) DUAT; 2) DUAT and AM weather; 3) FSS briefing; 4) FSS briefing and AM weather. Ten subjects were exposed to each type of weather format, and subjects were paid $5 for participating.

Materials. The materials consisted of a demographic questionnaire, a weather scenario, a flight plan questionnaire, and a sectional aeronautical chart.

Demographic Questionnaire. The demographic questionnaire consisted of 16 questions, including age, gender, type of pilot certification, number of flight hours accumulated, geographical area of the majority of flying, type of aircraft usually flown and type of method usually used to obtain a pre-flight weather briefing.

Weather Scenario. The weather scenario was collected on April 17, 1991 using equipment in the meteorology lab at ERAU during the time of 7:30 to 8:30 am. The AM weather television show, a verbal/map depiction of weather in a television presentation, was videotaped. A DUAT briefing, which is an encoded textual briefing, was obtained using a personal computer and modem, and was subsequently printed. Additionally, a call to a local FSS briefer, who gave a verbal briefing by telephone, was taped. The weather on this day was judged by several experienced pilots to be marginal for central and northern Florida. The critical weather problem was a low pressure area with thunderstorms in northern and central Florida moving east at 20 knots, with the thunderstorm cells within this area moving north-east at 15 knots. The weather was somewhat unusual for this time of year, as the low pressure area was causing the thunderstorms to be formed early in the day. The thunderstorms were not just a result of normal convective activity in the afternoon, as is more typical of Florida during the spring.

Flight plan questionnaire. The flight plan questionnaire consisted of eight open-ended questions regarding the pilots' decisions about flying in the given weather scenario. The

questions regarding the pilot's decision about time of day for departure, route and altitude for flight, and an alternate airport for landing. Each of these questions was followed by a question about the pilot's reasons for the previous answer.

Sectional Aeronautical Chart. The Jacksonville sectional aeronautical chart (National Oceanic and Atmospheric Administration, 1991) was also used. The chart presented a map of the central and northern sections of Florida, with prescribed airways, restricted areas and airport information depicted.

Procedure. Subjects were read the following set-up information for the weather scenario.

> "Today is April 17th. It is 8:00 am. You will be making a hypothetical flight from Daytona Beach to Tallahassee. You must fly sometime today. You will be flying an IFR-equipped, but not radar-equipped, CESSNA 172. You have a full fuel load, and will be carrying no passengers."

Subjects then studied the weather scenario in one of the four formats (DUAT, DUAT + AM Weather, FSS, or FSS + AM Weather) for up to 30 minutes. During this time the subjects could also use the Jacksonville sectional aeronautical chart to plan their flight. When the subjects were finished, or when 30 minutes had elapsed, the weather scenario and charts were collected. Subjects then completed the flight plan questionnaire using as much time as desired. Subjects were then paid and debriefed.

Results

How do the flight plan decisions of the subject pilots compare to the flight plan decisions of the expert pilots?

Five instructors from the Flight Line Department at Embry-Riddle Aeronautical University were convened as a group of expert instrument-rated pilots. They had a mean of 7320.00 total flight hours ($s = 2477.30$) and a mean of 595.00 flight hours of instrument rated flying time ($s = 519.74$). The expert pilots were asked to determine, in their best collective judgment, the optimal answers, or "best case" for the flight plan questionnaire. They were also asked to determine the least optimal answers, or "worst case" flight plan. The answers from the subject pilots were then scored twice, once using the optimal answers of the expert pilots as the correct response, and then a second time, using the least optimal answers of the expert pilots as the correct response.

For the optimal answers, subjects received one point for every answer that matched an expert answer, for a total possible of 8 points (based on route, time, altitude, alternate and reasons for each choice). The overall mean of the subject pilots ($M = 1.65$, $s = 1.23$) was significantly lower than the total possible, $t(39) = 32.65$, $p < .01$, showing that for the most part, the subject pilots did not choose the route, time, altitude or alternate airport chosen by the expert pilots. For the least optimal answers, subjects received one point for every answer that matched an expert answer, for a total possible of 4 points. The why questions were not

scored in this analysis, since the subject pilots were not giving reasons for choosing the "worst" flight plan. The overall mean of the subject pilots ($M = 1.75$, $s = 1.04$) was significantly lower than the total possible, $t(39) = 17.27$, $p < .01$. Thus, the subject pilots did not choose what experts consider to be the optimal flight plan, but they also did not choose the worst possible flight plan. This indicates that experience is an important contributor to understanding the weather and making appropriate flight decisions.

How does the weather information format influence the flight plan decisions made by the subject pilots?

Even though the flight plan decisions of the subject pilots did not match the experts, the decisions can be compared to each other to determine the influence of the four weather formats used in this study. The flight plan decisions for route, departure time, altitude and alternate airports will be considered separately.

The possible routes were categorized into three routes; a direct route – flying from Daytona Beach, to Gainsville, to Cross City, to Tallahassee; a northern route – flying from Daytona Beach north up the coast of Florida to Jacksonville, then heading west to Tallahassee; and a southern route – flying from Daytona Beach south to Orlando, then west to the Gulf coast of Florida, then north-west to Tallahassee. As reflected by Table 1, the majority of the pilots chose the direct route, which was also the route chosen by the expert pilots. Thus, although the subject pilots did not match the experts in the overall flight plan, they did tend to match the experts in the chosen route. A chi-square analysis indicated a marginal effect for type of weather format, $\chi^2 (6, N=37) = 11.39$, $p = .08$. All of the subjects receiving the DUAT + AM weather chose the direct route, as did the majority of pilots receiving the DUAT and FSS weather. The pilots receiving the FSS + AM weather, however, chose about equally between the direct route and the northern route.

Table 1. Frequency of three routes chosen by pilots receiving DUAT, FSS, and AM Weather Presentations.

	Weather Presentation Format				
Route	DUAT	+ AM	FSS	FSS + AM	TOTAL
Direct	6	9	6	4	25
Northern	4	0	1	4	9
Southern	0	0	2	1	3
TOTAL	10	9	9	9	37

The possible altitudes were categorized into five levels. The majority of the pilots were fairly evenly distributed into three altitude levels: 4000-4999 ($N = 10$), 5000-6999 ($N = 13$), and 7000-8999 ($N = 11$). The expert pilots chose an altitude of 7000-8000 feet. A chi-square analysis indicated no effect for type of weather format on choice of altitude, $\chi^2 (12, N = 40) = 7.73$, n.s.

The departures [illegible] gorized into six times. The most common choices were between 7 and 9 am ($N = 10$ [illegible] en 10 and 11 am ($N = 10$), and between 3 and 5 PM ($N = 9$). The

expert pilots chose a departure time between 3 and 4 PM. A chi-square analysis indicated no effect for type of weather format on time of departure, χ^2 (15, $\underline{N}$ = 39) = 11.76, n.s.

The possible alternate airports were categorized into six geographical areas. The majority of the pilots chose an alternate airport along the direct route chosen by most pilots. The expert pilots chose Panama City, which is west of Tallahassee. A chi-square analysis indicated no effect for type of weather format on alternate airport, χ^2 (15, $\underline{N}$ = 37) = 18.44, n.s.

How does the weather information format influence the reasons for the flight plan decisions made by the subject pilots?

This analysis looked at the why questions on the flight plan: or the reasons pilots gave for choosing a particular route, time, altitude and alternate airport. As above, each of these decisions will be examined separately. The reasons for each decision were categorized into three or four categories in consultation with the expert pilots.

The possible reasons for the route chosen were categorized into four reasons; time enroute, weather avoidance or concerns, navigational aids, and airspace restrictions or concerns. A chi-square indicated an effect for type of weather format on the reasons given for the chosen route, χ^2 (9, $\underline{N}$ = 40) = 19.46, $\underline{p}$ < .05. As can be seen in Table 2, the two FSS groups gave primarily weather related reasons, while the two DUAT groups gave more varied reasons. The two DUAT groups were combined and the two FSS groups were combined for follow-up comparisons. Orthogonal contrasts indicated that the differences between the FSS and DUAT groups lie in the weather related reasons as compared to all other reasons given, χ^2 (1, $\underline{N}$ = 40) = 10.16, $\underline{p}$ < .01.

Table 2. Frequency of four reasons for chosen route given by pilots receiving DUAT, FSS and AM Weather Presentations.

	Weather Presentation Format				
Reason for Route	DUAT	DUAT + AM	FSS	FSS + AM	TOTAL
Weather concerns	5	5	10	9	29
Time Enroute	1	4	0	0	5
Navigational aids	1	0	0	1	2
Airspace concerns	3	1	0	0	4
TOTAL	10	10	10	10	40

The possible reasons for the altitude chosen were categorized into four categories: navigational rules, safety, weather concerns, and performance of aircraft. A chi-square analysis indicated no effect for type of weather format on reasons given for the chosen altitude, χ^2 (9, $\underline{N}$ = 40) = 8.00, n.s.

The possible reasons for the time chosen for departure were categorized into four reasons: fly before the weather hazards are a factor, fly after the weather hazards have passed by, ceilings/visibility, and low turbulence/comfort. A chi-square analysis indicated no effect for type of weather format on reasons given for time of departure, χ^2 (9, $\underline{N}$ = 40) = 5.56, n.s.

The possible reasons for the alternate airport chosen were categorized into three reasons: weather concerns, time/distance, and navigational aids or other services. A chi-square analysis indicated no effect for type of weather format on reasons given for the chosen alternate airport, χ^2 (9, $\underline{N}$ = 40) = 10.13, n.s.

Many of these results presented for the flight plan decisions and the reasons given for the decisions are non-significant. An important finding, however, is that the choice of route seems to indicate an effect of the format of weather presentation. The graphical presentation (AM Weather) influenced the pilots who received an FSS briefing towards choosing indirect routes, while pilots who received the DUAT briefing and the graphical weather presentation tended to choose the direct route. This marginal effect was substantiated by the differences in FSS groups and DUAT groups for the reasons given for the chosen route; FSS groups gave more weather related reasons than DUAT groups.

Discussion

The study presented here is informative regarding methodological issues, however, it is only a first step in determining the effect of weather presentation format on situation awareness and decision making by pilots. Future studies should lead to making recommendations of an optimal presentation format for aviation weather and specific training techniques for pilots. The results of this study were somewhat surprising in that the type of weather presentation format appeared to have little effect on the types of flight decisions that were made by the instrument-rated pilots. It had been expected that the two groups of pilots receiving the graphical AM weather would make different decisions regarding each of the flight plan decisions, route, departure time, altitude and alternate airport, than the two groups of pilots who did not receive the AM weather. Although some differences did emerge for route, some other differences may have been masked by the variability of acceptable or "right" decisions. That is, there were a number of ways of conducting the flight safely according to the experts, thus the statistical power was reduced.

Although many of the results of the preliminary study were statistically non-significant, much was learned from conducting this research, and from subsequent discussions with researchers at NCAR. Importantly, we learned that the dependent variable that showed the most effect of presentation format was the choice of route. Thus, in follow-up studies, we will focus on using route choice as a measure of how weather presentation format affects decision making. Second, there was a lack of control in how the pilots' responses were assessed. Open-ended responses led to difficulty in scoring, and the interdependence between the questions on the flight plan questionnaire could not be taken into account statistically. That is, the choice of route is dependent upon the departure time, and this dependence could not be analyzed statistically. Third, it was noted that there was much variability between the pilots in their responses. This undoubtedly also contributed to the lack of significant differences. However, all of these problems were methodological in nature, and are being remedied.

We are currently working on developing a measurement of pilots' awareness of the current weather patterns, their ability to predict weather patterns, and their ability to determine how those weather patterns will impact flight decisions. Additionally, in an effort to reduce the

effect of individual differences in order to bring out any effect of weather presentation formats, we plan to use a within-subjects design.

The measurement we are developing is an objective questionnaire with clearly defined "better" and "worse" answers. We are currently focusing on thunderstorms, but expect that this methodology can be extended to a variety of types of weather. In the questionnaire, each question will present a departure time, and then several route choices will be listed as possible answers for that departure time. The route choices are defined as "better" and "worse" by objective criteria of severity of weather along the route. In this case, we are using the echo level through which the route passes. Thus, if pilots choose a route that takes them through a level 3 echo, it is objectively a "worse" decision than if they choose a route that takes them through a level 1 echo. Thus, pilots will be presented with weather scenarios in different formats and then asked to make these decision about routes for a variety of times for each weather scenario. We expect that when pilots receive the graphical format they will have more accurate awareness of the weather situation and will make better flight decisions than when they do not receive the graphical format. In addition to this research with flight-length weather scenarios and general aviation pilots, research is also needed to address the specific types of graphical formats for weather presentations that differ in their immediacy and severity of the weather hazards.

References

Bolstad, C. (1991). Individual pilot differences related to situation awareness. Proceedings of the Human Factors Society 35th Annual Meeting, pp. 52-56.

Cohen, G. (1989). *Memory in the real world.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Kit of factor-referenced cognitive tests.* Educational Testing Service: Princeton, NJ.

Endsley, M. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society, 32nd Annual Meeting* (pp. 97-101).

Endsley, M. (1990). Predictive utility of an objective measure of situation awareness. *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 41-45).

Garland, D. J., Stein, E. S., Blanchard, J. W., & Wise, J. A. (1992). Situational awareness in the future air traffic control environment. *Proceedings of the 37th Air Traffic Control Association Conference.* Atlantic City, NJ.

George, F. (1990). DUAT on-line: All the weather that's fit to print. *Flying, 6* pp. 81-83.

Glenberg, A. M. & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language, 31,* 129-151.

Glenberg, A. M. & McDaniel, M. A. (1992). Mental models, pictures, and text: Integration of spatial and verbal information. *Memory and Cognition, 20,* 458-460.

Levin, J. R. (1989). A transfer-appropriate-processing perspective of pictures in prose. In H. Mandl & J. R. Levin (Eds.), *Advances in psychology (Vol. 58): Knowledge acquisition from text and pictures,* 83-100. NY: Elsevier Science Publishers.

Mandl, H. & Levin, J. R. (Eds.). *Advances in psychology (Vol. 58): Knowledge acquisition from text and pictures.* NY: Elsevier Science Publishers.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82,* 715-726.

National Oceanic and Atmospheric Administration. (1991, March 7). *Jacksonville sectional aeronautical chart.* Washington, DC: U.S. Department of Commerce.

National Oceanic and Atmospheric Administration. (1991, March 7). *World Aeronautical Chart (CH-25).* Washington, DC: U.S. Department of Commerce.

Sarter, N. B. & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology, 1,* 45-57.

Schuyler, N. (1991, September). DUAT computer briefings are the only way to fly. *Private Pilot,* pp. 29-35.

Tenney, Y. J., Adams, M. J., Pew, R. W., Huggins, W. F., & Rogers, W. H. (1992). *A principled approach to the measurement of situation awareness in commercial aviation.* (Contract No. 4451). Washington, D.C.: NASA.

Wanke, C., Chandra, D., Hansman, R. J., & Bussolari, S. R. (1990). A comparison of voice and datalink for ATC amendments and hazardous wind shear alerts. *Fourth International Symposium on Aviation and Space Safety.* Toulouse, France.

Wanke, C., & Hansman, R. J. (1992). Hazard evaluation and operational cockpit display of ground-measured windshear data. *Journal of Aircraft, 29,* 319-325.

Author Notes

This research was supported by internal funds from the Center for Applied Human Factors in Aviation, University of Central Florida, and by a grant from the Division of Sponsored Research, University of Central Florida. The authors thank John Muller, and Jada Kearns for their assistance in conducting the research in this paper. In addition, the faculty at Embry-Riddle Aeronautical University, including, but not limited to, Tom Connelly, Frank Wencel, Jim Blanchard, Tom Kirton, Steve Hampton, Bill Kohrus, John Stratechuk and Joe Clark assisted in numerous ways. Also thanks to Wayne Sand and Eric Gruber for their help.

Aerospace Operations

- **Situational Awareness In Aerospace Operations: Personnel Training**
- **Key Situation Awareness Factors for Orbiter Flight Crew and Pad Ground Crew Emergency Egress and Escape Decisions During Space Shuttle Terminal Countdown Operations**

Situational Awareness In Aerospace Operations: Personnel Training

David L. Hosley

Lockheed Space Operations Company

Introduction

Lockheed Space Operations Company (LSOC) is the prime contractor for the National Aeronautics and Space Administration (NASA) Shuttle Processing Contract (SPC) to process the launching of the United States space shuttles at Kennedy Space Center, Florida. LSOC is assisted in this effort by other team member contractors: Grumman, Thiokol, and Johnson Controls. The combined workforce is approximately 8,000 employees. This paper describes the training provided for these employees so that they can effectively accomplish their mission.

Training

Personnel selected to work at Kennedy Space Center may not possess directly related experience. SPC Technical Training provides this knowledge and information so the employee can become productive in the shortest practical time. The work force of approximately 8,000 is divided into technical and administrative/other (Figure 1).

There is a training sequence for all new employees. Personnel hired by the SPC must meet the criteria in Company job descriptions for education and basic skills. Additional training on Shuttle Transportation System (STS) systems and skills may be required for specific job assignments.

Applicants for entry level positions or current employees seeking promotion to shuttle electrical or mechanical technicians and inspectors must pass a pre-employment qualification examination. The qualification examination is administered by the Vocational Department of Brevard Community College, Cocoa, Florida. Applicants who have completed requirements for a Federal Aviation Administration (FAA) Airframe and Mechanic's or Airframe and Power Plant Mechanic's license within five years of the date of application for employment will be exempt from pre-employment testing.

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

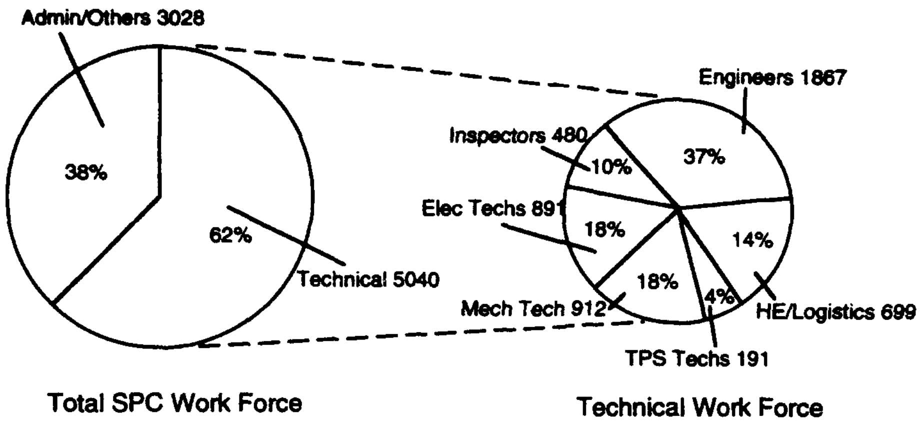


Figure 1. SPC Work Force (LSOC, GTSI, THI and JCWSI).

Orientation, Area Access, and Security

All employees must participate in an orientation, area access, and security program. The orientation part provides Shuttle Processing Contract (SPC) employees with an overview of the SPC organization, the objectives of the SPC program, Quality Assurance and the importance of Safety. This new employee orientation, conducted by Human Resources, includes such diversified subjects as employee benefits, security badge requirements, time cards, and KSC operations. The information is also reinforced by the employee's supervisor after reporting to the work area.

Area access requires that all employees will participate in general safety orientation that covers industrial safety, accident prevention, and fire protection. Employees receive formal credit for attending video walkdowns. Facility/safety walkdowns will be conducted by the employee's supervisor within three working days after assignment. The area access/safety training program is conducted by SPC Technical Training, Base Operations Contractor (BOC), SPC Safety and station managers/supervisors. Safety is recognized as a key element in all technical and operational training conducted by SPC. Refresher training is conducted on a recurring basis to maintain awareness. All employees participate in a general security orientation that covers security awareness at KSC. Refresher training will be conducted on a recurring basis to maintain awareness. This training is the responsibility of the BOC at KSC. Additional security training is provided by SPC as required.

Safety/Skills Training

Formal training classes in these two categories are initiated as soon as the individual's need is identified by supervision. Enrollment of the student is accomplished through their Training Coordinator who is located within the functional area.

Formal courses in safety are presented by SPC Technical Training, SPC Safety, and BOC Training. These courses provide SPC personnel with knowledge of safety hazards associated

with work areas. Specific courses are available from SPC Training on hazardous systems, i.e., pyrotechnics, cryogenic handling, high pressure, propellants, working at heights, equipment, etc. The BOC offers courses in Fire Safety and Medical training which are scheduled through SPC Technical Training. BOC Medical provides medical examinations for personnel whose job assignments could be affected by their physical condition, i.e., working at heights, heavy equipment operation, crane operations, soldering, SCAPE, and pyrotechnic handling. SPC attendance at BOC courses and physicals are monitored by the SPC Certification Board. Appropriate management is notified by the Certification Board Chairman when attendance is less than satisfactory. Skills courses are established to provide both theory and laboratory elements required by NASA, Launch Services Support Contractor's (LSSC's), and SPC management. The emphasis of this training is placed on the individual's ability to perform a specific skill (e.g., soldering, welding, non-destructive testing, bonding, heavy equipment operation, and crimping).

Task/Systems/Maintenance Training

Task and systems training is applicable to both Shuttle Transportation System (STS) processing operations and processing support functions. Formal training focuses on operation, maintenance, paper processing and paper closure for flight hardware, ground support equipment, and facilities. Maintenance training covers the application and operation of Bench Maintenance Equipment (BME); Pre-Installation Test (PIT) requirements, methods and procedures; and disassembly, assembly and check-out of complex or sophisticated equipment. It also includes courses designed to support unique refurbishment tasks. Launch processing System Integrated Ground Operations Maintenance training provides in-depth training for maintenance required on the Checkout Control and Monitor System (CCMS). Maintenance training covers development of hardware systems skills throughout CCMS and Central Data Systems.

On-The-Job (OJT) Training

The OJT Program is used to develop and maintain hands-on skills. The program uses OJT packages and requires proficiency demonstrations to ensure the individual has achieved the necessary level of performance. The roles and responsibilities for functional organizations and the SPC Technical Training Department are provided in Company operating procedures.

Qualification/Certification

Certain processing and inspection tasks are performed by personnel who are required and/or certified to perform these functions. An individual is considered a candidate for certification only whe he has satisfied all the requirements imposed by the Certification

Board for the particular task or skill. After an initial certification, Training and Certification Record System (TCRS) reports provide information to department training coordinators so that recertificauon (retraining, proficiency testing, and/or physical examinations) may be scheduled to renew current certifications prior to their expiration.

Cross Training

Each functional organization is responsible for identifying personnel assigned to cross training and for scheduling additional class room training as required. Personnel selected for cross training will have been previously qualified and/or certified on other job family skills.

Figure 1 portrayed the work force of approximately 8,000. As the above information indicated, there is a lot of training being conducted. These requirements are driven by the need for a highly qualified and certified workforce. Table 1 provides a snapshot of t he FY 92 monthly averages and the cumulative data for certification and courses.

SPC Technical Training does not conduct flight training or flight simulation. The training does stress quality, safety and an overall awareness to the working environment in each of the training courses. Training contributes to the operation of the shuttle in space, and it plays a vital role in each of the 1,000,000 tasks that are performed for each launch.

Table 1. Shuttle Processing Contract (SPC) Training Environment

FY 92 Technical Training Elements

Monthly Averages	
Initial Requirements	4,635
Classes Conducted	373
Students Attending	2,685
Student Hours	10,172
Data	
SPC Certification Criteria Sheets	535
Average Certification per Employee	7.1
SPC Employees with Certifications	4,227
On-The-Job Training Packages	579
SPC Training Courses	382
Video Tapes	188

Situational Awareness

The Technical Training environment does not use the situational awareness definitions that are usually affiliated with a flight cockpit and its changing internal and external

environment. Situational awareness to the SPC workers is relative to the changing position of the shuttle during processing and its effect on the tasks to be performed. The movement from the horizontal to vertical position of the shuttle as it moves from one processing area to another, culminating with its location at the launch pad, places an utmost need for workers to be aware of their environment. The flight hardware is unique and very expensive. There are some 24,000 tiles and 9,200 blankets that are a part of the thermal protection system to protect the orbiter inside and out from the searing heat of launch exit and re-entry – as high as 3,000 degrees Fahrenheit. Special switches, electrical connectors, fittings, etc., throughout the orbiter cost thousands of dollars each and must function perfectly in the isolation of space. Each worker must be extremely sensitive and attentive to detail in the performance of work. The horizontal postflight servicing, checkout and modifications performed in the Orbiter Processing Facility (OPF) provide an entirely different orientation than the vertical processing at the Vehicle Assembly Building (VAB) and Pads.

Examples Of Courses

Table 1 included the number of SPC training courses (i.e., 382) that are conducted at KSC for the SPC. These formal courses are in the subject areas described earlier in this paper. The following examples best fit our situational awareness environment.

QS-22B-LSC, Flight Crew Emergency/Egress Escape and Rescue (walkdown)

This 2-hour walkdown course is to familiarize the Orbiter Flight Crew, Closeout Crew, and Pad Rescue Team with procedures, routes, and equipment used during an Emergency Egress/Escape/Rescue. This includes STS Contingency Management, STS Contingency Forces, Fallback Areas, Pad Ingress/Egress Routes, Fixed Service Structure (FSS) Level 195' Overview, Crew Egress/Extraction, Slidewire System, Bunker/Pad Egress, Emergency Medical Services Site, helicopter Familiarization, Nighttime Egress, and a walk through simulation. (Before launch and 12 months recurring period.)

OC-200-LSC, Payload Flight Crew Systems

This 16-hour course is designed to familiarize technicians and inspectors with flight crew systems orbiter processing facilities (OPF) procedures that configure both flight and general support hardware (GSE), as well as required testing procedures. (Not recurring.)

QG-245-LSC, Working at Heights Safety

This 2.5-hour course is to familiarize the student with fall protection equipment and temporary access devices encountered at Kennedy Space Center. The proper inspection, donning, and use of personal fall protection equipment is thoroughly discussed. (36 months recurring period.)

OV-599-LSC, Monoball Installation

This 8-hour course plus an on-the-job requirement is designed for the technician to demonstrate the proper assembly/disassembly of the monoball connector and the proper procedures for installation of the monoball assemblies. (24 months recurring period.)

OV-257-LSC, Orbiter Environmental Control and Life Support System (ECLSS) Familiarization

This 3-hour course provides general information on the ECLSS. It is intended for personnel who need to understand associated subsystems and their primary functions. The interface of ECLSS with other systems is thoroughly discussed. (No recurring period.)

Conclusion

The space shuttle program is the world's best. It demands the highest levels of quality and safety. The SPC Technical Training program offers a wide variety of courses – 382 formal courses, 579 OJT packages, and 188 videotapes – to meet these training needs. Each and every person must be sensitive and aware of their working situation and environment. The training program accomplishes this.

Key Situation Awareness Factors for Orbiter Flight Crew and Pad Ground Crew Emergency Egress and Escape Decisions During Space Shuttle Terminal Countdown Operations

Eric F. Redding

National Aeronautics and Space Administration

Introduction and Background

During the Space Shuttle terminal countdown phase there are numerous failures within the flight hardware and the associated ground support equipment that would necessitate an emergency egress and escape by any personnel at the launch pad. The terminal countdown phase corresponds to the start of external tank cryogenic loading through launch or launch abort safing operations. These failures can be categorized into specific criteria as listed below [1]:

- Fire external to any vehicle element (Orbiter, External Tank, Solid Rocket Motors), which is uncontrollable and jeopardizes the External Tank or Solid Rocket Motors.
- Fire internal to any vehicle element.
- Ruptured fuel or oxidizer line inside any vehicle element that could lead to an explosion or fire.
- Uncontrolled pressurization of tanks or vessels on or inside any vehicle element.
- Fire or hypergolic leak (monomethyl hydrazine, nitrogen tetroxide, or hydrazine) external to any vehicle element that would prevent normal egress by the flight crew.
- Loss of gaseous nitrogen purge to the Orbiter aft and midbody compartments and liquid oxygen or hydrogen leakage in the Orbiter aft compartment or the Tail Service Mast.
- Hydrogen leakage in the Orbiter aft compartment in excess of one percent (10,000 parts per million).
- Loss of Orbiter power and the resulting loss of control and monitoring capability of the Space Shuttle systems.

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

It is during this phase that the following launch team personnel could be on the launch pad itself and it is the focus of any launch pad egress and escape decision to assure the safety of these personnel.

Flight Crew – Five to seven astronauts.

Closeout Crew – Six or seven individuals (depending upon Flight Crew size) who enter the pad after External Tank cryogenic loading to conduct planned operations to prepare the Orbiter for Flight Crew ingress, to conduct Orbiter crew module hatch leak check, and to close out the crew module hatch and access room for launch.

Pad Fire and Rescue Team(s) – Consists of two contingency teams. The Pad Fire and Rescue Team has 8 personnel from the Kennedy Space Center Fire and Medical services. The second team is called the Pad Fire and Rescue Driver team which consists of two personnel with the prime responsibility of operating a pre-staged armored personnel carrier used to evacuate personnel from the pad area to predesignated helicopter pickup points and triage sites. Prime responsibilities include aiding and assisting pad personnel who are injured or incapacitated and conducting fire fighting operations.

Ice/Frost Inspection Team – Eight personnel who are scheduled to enter the launch pad after External Tank cryogenic loading and prior to Flight Crew departure for the pad to conduct visual observations and obtain data on the Shuttle and launch complex equipment and facilities based upon the cryogenic and atmospheric environment.

Red Crew – Contingency crew of engineers and technicians dispatched to the pad to troubleshoot and repair pad hardware that is deemed critical for a successful and safe launch.

Any Red Crew or Ice/Frost Inspection Team pad activities should be finished prior to sending the Flight Crew to the pad. Because of this planning, the Red Crew or the Ice/Frost Inspection team are not considered in this particular discussion on emergency egress or escapes.

Modes of Emergency Egress/Escape

There are four specific modes of emergency egress options available to the above pad personnel. These modes are primarily differentiated by which of the above personnel are on the pad at particular time and/or who is needed to effect an escape [2]:

Mode 1 - Condition when the Flight Crew is in the Orbiter crew compartment and is able to egress and escape without assistance.

Mode 2 - An aided emergency egress and escape of the Flight Crew performed by the Closeout Crew.

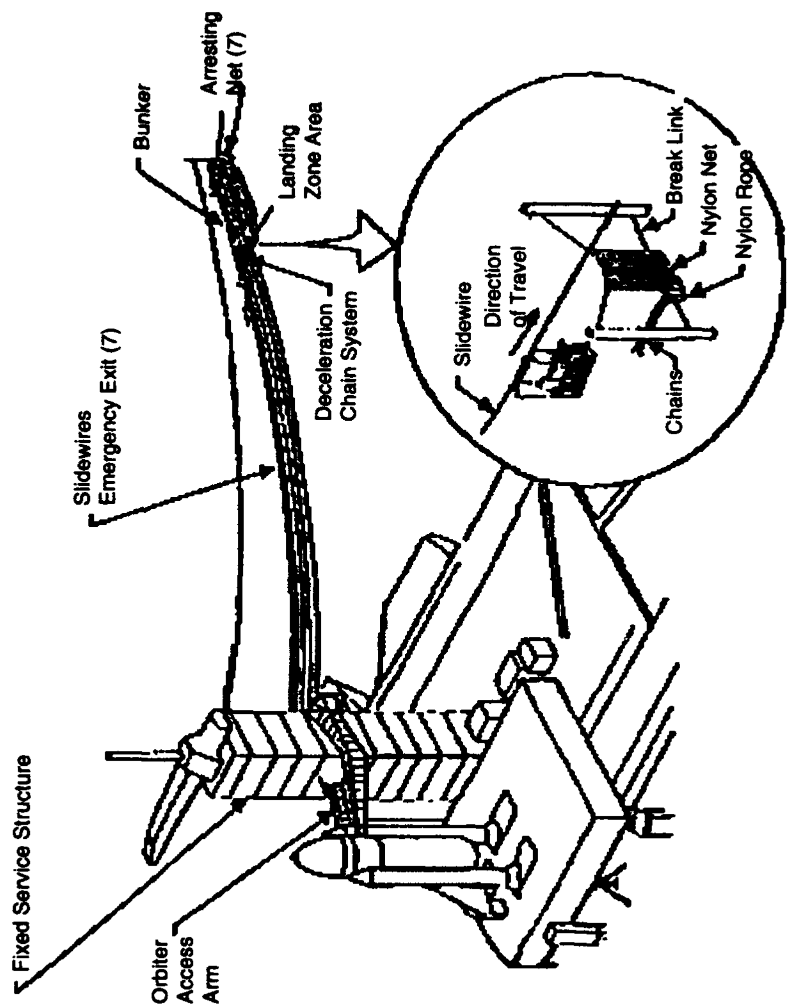

Fixed Service Structure
Orbiter Access Arm
Slidewires Emergency Exit (7)
Deceleration Chain System
Bunker
Arresting Net (7)
Landing Zone Area
Slidewire
Direction of Travel
Break Link
Nylon Net
Nylon Rope
Chains

Figure 1.

Mode 3 - An emergency egress condition in which the Closeout Crew is not on station, and the Pad Fire/Rescue Team is required to proceed from a fallback area to the Pad Orbiter Access Arm area and rescue an incapacitated Flight Crew.

Mode 4 - An aided Emergency egress and escape of the Flight Crew performed by the Pad Fire/Rescue Teams while the Closeout Crew is on site at the Orbiter Access Arm or crew module.

Pad Escape Systems

The pad Emergency Egress and Escape System is the primary means of egress and escape for those pad personnel identified previously. The system utilizes seven separate slidewires with seven multi-man basket assemblies to effect a rapid escape of personnel from the Orbiter Access Arm level on the Fixed Service Structure (195 foot level) to a landing zone approximately 1200 feet west of the pad structure to where the blast protection Bunker is located. Each slidewire basket is designed to accommodate three personnel and is routinely validated to operate safely with a four person load. The nominal capability of the system as a whole is twenty-one personnel. (FIGURES 1 and 2) [3]

The blast bunker is an earth-covered, steel-reinforced concrete bunker that provides a safe haven for pad personnel in the event of a pad emergency in which the slidewire system is used. It contains first aid equipment, instrumentation to detect toxic fuel or oxidizer vapors and flammable conditions, breathing air supply and distribution equipment, and a direct communications link to the NASA Test Director via a point-to-point phone. A second M113 armored personnel carrier is parked outside the bunker for pad egress and evacuation.

Monitoring Systems

There are approximately 2500 measurements that are remotely monitored by the Launch Processing Computers and the launch team. Many of these measurements are used to monitor and determine the health and safety of the Flight Crew, the launch vehicle and the associated ground support equipment. Measurements such as pressures, flow rates, temperatures, fire and smoke detectors, leak concentrations, and many others are used to detect and evaluate failures such as the those specified emergency egress criteria discussed earlier. In addition to these measurements, launch team personnel also rely on up to eighty-three remotely controlled television cameras strategically positioned on the launch pad, as well as on-site observations made by the Flight Crew, Closeout Crew, and the Fire and Rescue Teams. These on-site observations are then transmitted verbally to the Launch Control Center and to the rest of the launch team over radio nets or on an Operational Intercommunications System. Other measurements are used to aid the launch team in reacting to such failures and effecting a rapid and safe emergency escape. Measurements such as temperatures, fire and smoke detectors, leak detectors, and oxygen deficiency monitors are used to assess the conditions of the

possible egress and escape options and aid in the implementation of a safe and rapid egress decision. Weather conditions are also monitored from remote instrumentation at the pad and surrounding areas. These decisions are also augmented by on-site and television monitor observations.

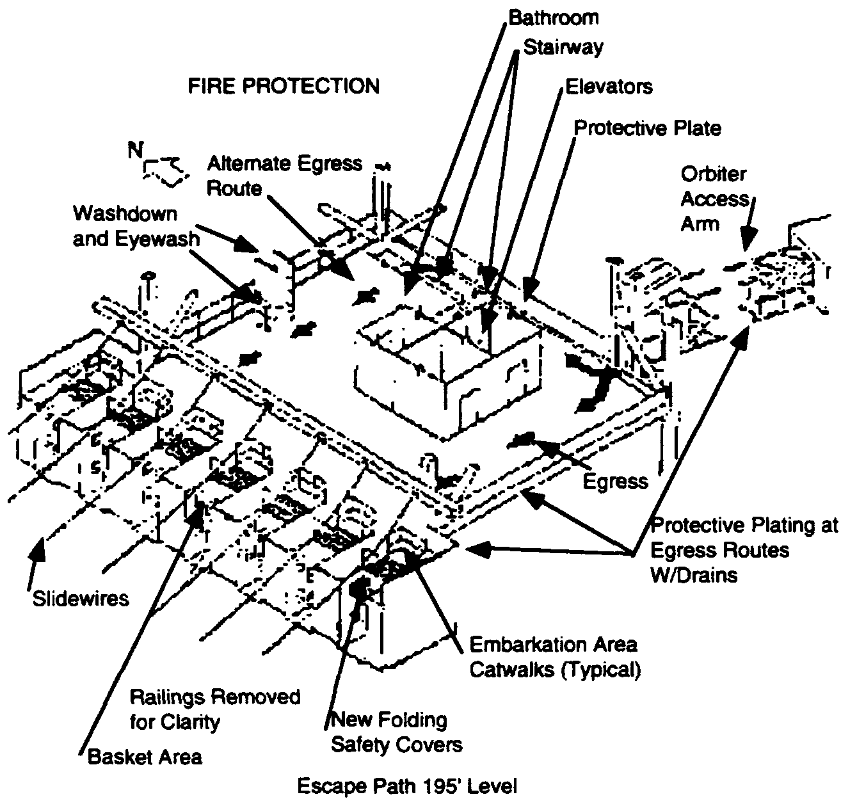


Figure 2. FSS Egress Route showing flame barriers

Responsibilities

There are four essential members of the launch team that have either the authority to initiate an emergency egress and escape and/or have very important responsibilities during an emergency egress and escape.

NASA Test Director. Senior launch team manager that is responsible for the tracking and accountability of all personnel within the launch pad, the Launch Danger Area, and the Blast Danger Area during the terminal count phase. This individual is responsible for making the

decision recommending an emergency egress to the Flight Crew and for the initiation and direction of the appropriate actions. He is also responsible for determining the appropriate response to a critical condition as identified by the other members of the launch team and will determine if immediate action is required to ensure the safety of the Flight Crew and pad personnel. He will also respond to an emergency egress and escape that is initiated by the Flight Crew Commander or Closeout Crew Leader.

Closeout Crew Leader. On-site contractor technician who leads crew module closeout activities and has responsibility for directing on-site emergency egress and escape activities of the Flight Crew during the Flight Crew insertion and closeout activities, and for decisions regarding safety of Closeout Crew during these emergency actions. The responsibility of the Closeout Crew Leader during an emergency egress and escape terminates with the delivery of the Flight Crew to designated medical personnel or upon leaving the elevator at the pad surface level and returning to their designated prelaunch staging area during a nominal launch.

Pad Fire and Rescue Leader. Senior Contractor Fire Department Officer who is responsible for implementing assigned mode of emergency egress and escape activities and for on-site decisions regarding the safety of Fire and Rescue personnel during these emergency operations.

Flight Crew Commander. Senior Flight Crew pilot with command authority for the Flight Crew. Is responsible for on-site decisions regarding safety of the Flight Crew during a MODE 1 egress.

Procedures

There are several NASA publications that have been developed to aid in the training for and execution of an emergency egress and escape at the launch pad.

S9913 – Orbiter Flight Crew and Pad Ground Crew Emergency Egress and Escape at the Launch Pad. This is a working level Emergency Procedure Document that is implemented by the NASA Test Director in an emergency situation that warrants such as egress and escape. It contains the detailed steps, diagrams, egress criteria, and available options that the NASA Test Director would implement in such an emergency.

KVT-PL-0004 – Space Transportation System Operations Plan for the Pad Emergency Egress System. This plan establishes the Kennedy Space Center policy for using the pad emergency egress system, provides a detailed description of the system and provides the plan for pad personnel accountability during an emergency egress. The plan ap[illegible]o and describes all potential users of the system and those personnel responsible [illegible] operation and maintenance. This plan also specifies policy for configuration control of the system.

S1025 – Flight Crew and Ground Crew Emergency Egress and Escape Test. Detailed test procedures used for periodic training exercises. Objective of these tests include familiarity of the Orbiter Flight Crew and Ground Crews with the available evacuation routes, emergency

equipment, and personnel that would be involved in a pad emergency egress and escape. These exercises are conducted for each mission as part of the Countdown Demonstration Test and occurs approximately three weeks prior to launch. These exercises are also augmented by several training classes conducted during this same time period.

S0044 – Shuttle Final Countdown Phase Simulation. Detailed training procedure for conducting simulation training on a complex math model. This exercise subjects the entire launch team to numerous failures that must be detected and resolved. This integrated test is performed twice each mission.

S0007 – Shuttle Final Launch Countdown. Detailed integrated procedure used to conduct the launch countdown. Volume 2 contains the control sequences while Volume 5 contains contingency procedures and emergency instructions.

NSTS 16007 – Shuttle Launch Commit Criteria and Background. Contains the detailed subsystem level criteria that must be met to proceed with launch. Also contains preplanned decisions that have been designed to minimize the amount of real time rationalization required when off-nominal situations occur.

Training

There are several training exercises and simulations conducted at Kennedy Space Center each Space Shuttle mission to prepare the Flight Crew and ground crews for pad emergencies. Detailed briefings and on-site pad walkdowns are given to the NASA Test Director, the Flight Crew, Closeout Crew, and the Pad Fire and Rescue Team one day prior to the Terminal Countdown Demonstration Test. The objectives of this test are to provide detailed audio and visual instructions coupled with hands-on training at the launch pad. This training also includes refresher courses in the use of several breathing air devices and Cardio Pulmonary Resuscitation. These activities also allow these key team members the opportunity to interface as a team and discuss face-to-face any concerns there may be.

A Terminal Countdown Demonstration Test is also conducted and has primary objectives of demonstrating and evaluating the Flight Crew operational time line, the interfaces between the Flight Crew and the Launch Team, and to demonstrate launch abort and recycle operations. The launch abort saving and recycle portion includes a MODE 1 egress simulation that allows the Flight Crew to practice what has been taught in the previous day's classroom training and walkdowns.

In addition to the above exercises, simulated launch countdowns and propellant loading exercises are also conducted on a large computer based math model for each mission. The primary objective of these exercises is to train and evaluate the systems engineers and launch management on the detection, reporting and resolution of inserted failures. Often failures are inserted that warrant an egress and escape decision.

In addition to the above required training, Pad personnel also receive periodic and mandatory classroom and hands-on instructions on the pad facility, its associated ground

support equipment, rescue and fire training, emergency equipment, and Orbiter subsystems in order to become launch certified.

Scenario

The countdown clock is at T-30 minutes and counting, with all five Flight Crew members secured in their ascent seats in the Orbiter crew module. The six member Closeout Crew is closing the Orbiter side hatch and securing the access room for launch. The external tank was previously loaded with 223,000 gallons of liquid oxygen and 466,000 gallons of liquid hydrogen and is now in a stable replenish mode. The Orbiter payload bay and aft compartment is being purged with gaseous nitrogen. The two-man M113 armored personnel carrier Driver Team is stationed within the pad perimeter in the blast protection bunker. The eight member Fire and Rescue Team is dressed out in their protective gear and is standing by at the Blast Danger Area roadblock, some 4485 feet from the pad center, ready to enter the pad area should they be called upon.

Suddenly Mission Specialist 3 reports a loud noise, possibly an explosion, in the payload bay area and that he has multiple injuries up on the flight deck. The Commander confirms the same and reports Mission Specialist 1 has a badly injured shoulder, Mission Specialist 2 's sight is impaired, and the left aft flight deck view port has been dislodged from its structure and gaseous nitrogen is flooding the crew module. Simultaneously several systems engineers begin reporting multiple systems failures, but the NASA Test Director quickly holds these calls and declares a MODE 2 Egress. As the Closeout Crew assures the Flight Crew is on breathing air and aids the crew in egressing the crew module, the NASA Test Director is directing the activation of Firex water flow on the Orbiter Access Arm and the adjoining level to combat any fires that may impede egress by the Flight Crew and Closeout Crew. He has also initiated the move of the Pad Fire and Rescue Team from the Blast Danger Area roadblock to the pad. Immediately following these actions, the Closeout Leader reports to the NASA Test Director that Mission Specialist 1 and 2 are incapacitated and additional assistance is needed. The NASA Test Director immediately declares a MODE 4 and notifies the Pad Fire and Rescue Team by radio of the call and to continue on to the pad, further instructions forthcoming. The NASA Test Director quickly ascertains the integrity of the launch vehicle and pad and is able to verify that there are no fuel or oxidizer leaks or fires or any other conditions that would preclude sending the Pad Fire and Rescue team into the pad . He advises the Pad Fire and Rescue Team of this and the need to wear breathing air equipment due to the gaseous nitrogen flowing into the crew module. He then commits the crew into proceeding into the pad perimeter and up the structure to aid in the rescue.

In the mean time the Closeout Crew has assisted in getting the Pilot, Commander, and Mission Specialist 3 into one slidewire basket and releasing it to the slidewire termination area. A second basket has also departed with three members of the Close-out Crew. In less than four minutes, the Pad Fire and rescue Team has proceeded up the structure and aided the remaining Closeout Crew in transporting the two downed Flight Crew members into the third slidewire basket and has released them along with one Closeout Crew member. The remaining two Closeout Crew members and the Pad Fire and Rescue Team enter the remaining baskets and are quickly away to the bunker area. It is during this time that the

Flight Crew Commander has made contact with the NASA Test Director on the point-to-point phone and a decision has been made to leave the bunker area via the two M113 armored personnel carriers and proceed to the optimum helipad for triage treatment and medevac to a previously selected and notified hospital.

Key Situation Awareness Factors

The time elapsed from problem detection through bunker area debarkation was less than seventeen minutes. In this small period of time several time, critical decisions were made to assure a rapid and safe egress and escape. Situation Awareness, on the part of the four essential Decision Makers, played a very crucial part in this scenario. But what factors, or characteristics, are present in this situation that helped maximize the level of Situation Awareness?

The foundations of good Situation Awareness begins with established procedures, policies and a thorough knowledge and experience of the tasks, equipment, and capabilities involved. These factors can be exercised and tested in various training exercises and simulations. But this is usually done in an artificial environment under ideal conditions. These conditions are rarely present in a real world emergency. The information flow from the various television monitors, verbal communications from pad personnel and launch team members, instrumentation, and various countdown clocks further increase the level of Situation Awareness. All of these factors have some limitations and is it because of these limitations that one often finds difficulty in applying this knowledge at the right place, at the right time, and in the right sequence, in such a dynamic and dangerous emergency situation. However, there are additional factors that can help one achieve an even higher level of Situation Awareness and it is these factors that the remainder of this paper will discuss.

The first factor that the Decision Maker, or DM, must have is a good concept of Time. Time, as discussed in this context, has three aspects. Time, in the first aspect, is defined as a function of the launch vehicle configuration and it denotes sequential events that occur in a launch countdown that change the launch vehicle or ground support equipment configuration such that different options are available to safe the vehicle and crew. An example of this is the issuance of the Space Shuttle Main Engine start commands at T-6.6 seconds and the different actions required to safe the vehicle and crew should a launch abort occur. Another Time aspect deals with time constraints and how long one has to get the necessary information, make an evaluation, and make and execute a decision without subjecting the crew or launch vehicle to any critical situation any longer than absolutely necessary. The third aspect of Time requires the understanding of how long each task takes, in what order they must occur, and how to optimally fit them all into the whole operation to achieve a safe and rapid egress and escape. This requires looking ahead and looking at the whole picture. An example of the above aspects was the NASA Test Director dispatching the Pad Fire and Rescue Team to the pad, knowing it would take them approximately four minutes to reach the pad gate, and updating this crew in route with the launch vehicle integrity and pad conditions prior to committing the team through the gate and up the structure.

Another factor that is critical to the DM in gaining an elevated level of Situation Awareness is Spatial Orientation. Spatial Orientation, in this context, depicts the physical

layout and capabilities of the pad structure, the Orbiter crew module, the Orbiter Access Arm and its adjoining level, the pad Emergency Egress System, and the locations of each individual on the pad. This factor also includes the tracking of these pad individuals and is often referred to as Personnel Accountability. This aspect is very important for the NASA Test Director, as it is his responsibility to know where each pad individual is at all times. This factor aids the DM in understanding the overall logistics of the emergency and helps ensure the proper aid gets to the right place and to the right individual. Development of this factor requires a very thorough knowledge of the pad and Orbiter layout. This knowledge must be maintained by frequent visits and walkdowns of these locations.

The third factor that will be discussed is Decision Authority. This factor addresses the rationale used to delineate decision making authority in an emergency situation. Usually this authority has been previously delegated by established procedures and policies. The delegation of this authority should be based on which individual has all the necessary information to make the optimum decision (i.e. the Big Picture). As in the case of a pad emergency egress and escape, the NASA Test Director has several sources of information to aid him in making the egress decisions and committing the Pad Fire and Rescue Team into the pad and up the structure. Failure to look at the whole situation and ascertain the vehicle and pad integrity prior to committing this crew would put additional personnel at risk. However, it must be pointed out that most of this information that the NASA Test Director obtains is artificial visual and verbal inputs from outside sources and can be subject to delays and misinterpretation.

Conclusions

By beginning with detailed policies and procedures, sound training programs, and a detailed knowledge of the hardware and its operation, one can develop the basic foundation for Situation Awareness in decision making. However, real time decision making in a real world emergency situation, such as that of a very complex and dynamic Space Shuttle Launch countdown, requires a heightened level of Situation Awareness. Increased visibility into the performance of the complex systems through such things as instrumentation, television monitors, and on-site personnel can increase this level of awareness. However, a very thorough understanding of key factors such as Time, coupled with good Spatial Orientation and the selection of a Decision Maker who has the "Big Picture" can help achieve the level of Situation Awareness necessary to effect a safe and rapid pad emergency egress and escape.

It is also very important to note that the tendency to rely on technology to provide increased Situation Awareness is all too common in today's environment. However, as realized in this paper, it is some very basic Human Factors and the Human element itself that play a very crucial role in an emergency management situation. For machines and devices do fail and it becomes necessary for the individual human to interpret the situation using his past experience and expertise and respond accordingly.

References

[1] S9913 Revision O (November 3 1992) - Orbiter Flight Crew and Pad Ground Crew Emergency Egress and Escape Launch Pad. (Emergency Procedures Document) Kennedy Space Center, Page 24.

[2] S1025 Revision f (April 15, 1992) - Flight Crew and Ground Crew Emergency Egress and Escape Test. Kennedy Space Center, Pages 25-27.

[3] KVT-PL-0004 STS Operations Plan for the Emergency Egress System at Launch Complex 39 (DRD #017) September 14, 1988. Page 50 Figure A-18, page 34 Figure A-2.

[4] S0007 Volume 2 Revision BE (December 9, 1992) - Shuttle Final Launch Countdown. Kennedy Space Center.

[5] S0007 Volume 5 Revision K (December 11, 1992) - Shuttle Final Countdown, Contingency and Emergency Procedures. Kennedy Space Center.

[6] NSTS 16007 - Shuttle Launch Commit Criteria and Background. Revision F, change No. 018 (January 11, 1993).

Emergency Management Systems

- Situational Awareness in Emergency Management Systems: An Overview
- Situational Awareness in Emergency Services Incident Command

Situational Awareness in Emergency Management Systems: An Overview

Michael A. Companion

University of Central Florida

Introduction

Emergency management encompasses a broad range of civilian applications including law enforcement, fire fighting, hazardous materials, and a variety of other emergency response situations, e.g., hurricanes, nuclear accidents, aircraft accidents, and earthquakes. The decisions made by today's emergency response personnel (law enforcement, fire fighters, paramedics and emergency medical technicians) can significantly impact the communities they serve. Generally, emergency response personnel are considered to have influence only while the situation is in progress. However, their decisions have direct repercussions on lives, the environment, and the economy of a community. The results may be felt for many years. This is true of day to day decision making as well as major incidents. Today's emergency services are involved in all aspects of disaster planning and response, including natural and human-made disasters and hazardous materials incidents, in addition to the day to day operations of the departments. These large scale incidents require even greater organizational, situational awareness, and decision making skills on the part of participants.

Examples of the impact of decisions made by emergency management personnel in real life situations abound. These decisions can have a significant impact on the economy and environment as shown by these examples.

- A fire chief properly elects to stop fighting a fire and let it burn because it stands on the aquifer for the town. A national firm's profit margin is dramatically affected for years to come because of this decision.

- A hazardous material incident devastates an internationally known river. The economic well-being of thousands of people in the region is jeopardized – their health is threatened. The river is destroyed for years.

- A lone police officer makes the wrong decision and the town is eventually thrown into bankruptcy because of litigation costs and awards.

The scope of emergency services management continues to expand. Greater demands in the face of financial and manpower reductions necessitate more creative utilization of

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

available training resources. Emergency situations encountered by law enforcement, fire departments, and other emergency management agencies and groups require significant tactical decision making skills to achieve effective and efficient resolutions to incidents. Improper or slow decision making and/or situation awareness can significantly increase loss of life, loss of property and damage to the environment. While many techniques are available to train physical skills, there are few effective training techniques and tools to teach critical decision making skills. The use of conventional techniques for decision making training are not efficient because of

- the cost to conduct large scale simulated emergencies
- the risk to the participants
- the limited number of persons who actually receive decision making training
- the unique nature of each type of emergency situation.

There is a need to understand the situational awareness and complex decision making skills associated with emergency response so that effective training techniques can be developed to acquire these skills (Kass et. al., 1991).

The Situational Awareness Problem

Situational awareness is one of the underlying fundamental complex human behaviors. It is the process by which a person extracts, integrates, assesses and responds to task relevant information from the total environment, both spatial and temporal. In many real world situations, situational awareness becomes a critical skill for survivability. However, the basis for this skill-based behavior is not well understood. The acquisition of situational awareness skills also tends to be developed on a trial and error basis. Widespread interest exists in understanding this generic problem for complex systems.

Complex operational systems are limited primarily by the human component. This is especially true for emergency management events. Situational awareness or the operator's ability to extract, integrate and assess task relevant information from the total environment is critical to the effectiveness of many complex systems. It reflects a key human limitation to cope with the complex and interactive set of variables in the total environment. To achieve improved effectiveness of complex operational systems, it becomes imperative to gain an understanding of the cognitive processes involved in situational awareness, and to develop methods or tools to augment the human's ability to acquire this necessary skill.

An examination of the literature reveals numerous psychological theories that provide insight on the problem of situational awareness. The most important aspect is the synergism of these diverse behavioral theories when linked with situational awareness. These theories all converge on the notion that situational awareness is a skill-based behavior involving complex dynamic pattern recognition.

Categories of Behavior

Human behavior can be organized into three basic hierarchical categories as postulated by Rasmussen (1986). These three categories of behavior, knowledge-based, rule-based and skill- based, differ in detail and level of complexity. Knowledge-based behavior represents the most basic form of behavior. Information in this level of behavior is generally not integrated and has to be processed individually. Rule-based behavior reflects the human's attempt to simplify the environment by integrating information. Skill-based behavior is the most complex, yet fastest, category of behavior. In skill-based behavior, the individual appears to be responding to complex patterns of stimuli or features within the environment which elicit automated patterns of response.

These classes of behavior are not independent. They represent progressive stages of cognitive processing which reflect the individual's degrees of learning or experience. Each level of processing results from an increasing degree of knowledge compilation and integration. Just as rule-based behavior derives from the integration of information and responses (i.e., knowledge), skill-based behavior derives from the proceduralization of rules. The automaticity of skills indicative of skill-based behavior, such as situational awareness, is an outgrowth of the process of knowledge compilation which is founded in knowledge-based and rule-based behaviors. The pattern recognition which underlies the skill-based behavior comes from some level of processing that relates or compiles information to form a pattern. The resulting pattern/response combination is simply a highly complex or "ultimate" rule. This systematic progression of knowledge compilation from rule-based to skill-based behavior indicates that it should be possible to develop training strategies to positively influence the acquisition of situational awareness skills (Companion et. al. 1990).

This underlying premise which appears consistent both in theory and experience is that pattern recognition is the essence of the skilled behavior, of situational awareness. The individual learns to respond to patterns of variables in the environment. However, it must be noted that situational awareness is also a temporal phenomenon. Because of the compression of time and demand for quick actions in a tactical environment, situational awareness must encompass not only the current real-time situation, but also include anticipatory awareness of the situation. Situational awareness is a temporally dynamic process.

Overall, the concept of situational awareness can be perceived as a dynamic pattern recognition process (Figure 1). It can be viewed as a response space with the three dimensions: the relevant situational variables, time and criticality. The criticality of a variable is learned and changes as function of time based on the situation. These factors effectively describe a response surface. People tend to initiate responses when a pattern of variables exceeds some criterion. The use of a response criterion is the human's innate way of managing workload and works on the principle that "if it isn't broke, don't fix it." The response criterion is not a fixed entity. It may change as a function of learning, physiological state, motivation, stress, etc.

If the response criterion is viewed as a plane cutting through the situational response surface, then, as shown in the insert, it can be illustrated as a two dimensional plane overlaid by the pattern of variables which exceed the response criterion. This is a constantly changing pattern due to the dynamic nature of the phenomenon. It is conjectured that these patterns elicit plans of action. Unlike rule-based models of cognitive behavior, which may have hundreds of active plans at any moment in time, our concept suggests that only one plan is active at any one time but that it changes dynamically. This later concept is both parsimonious and correlated to observed behavior.

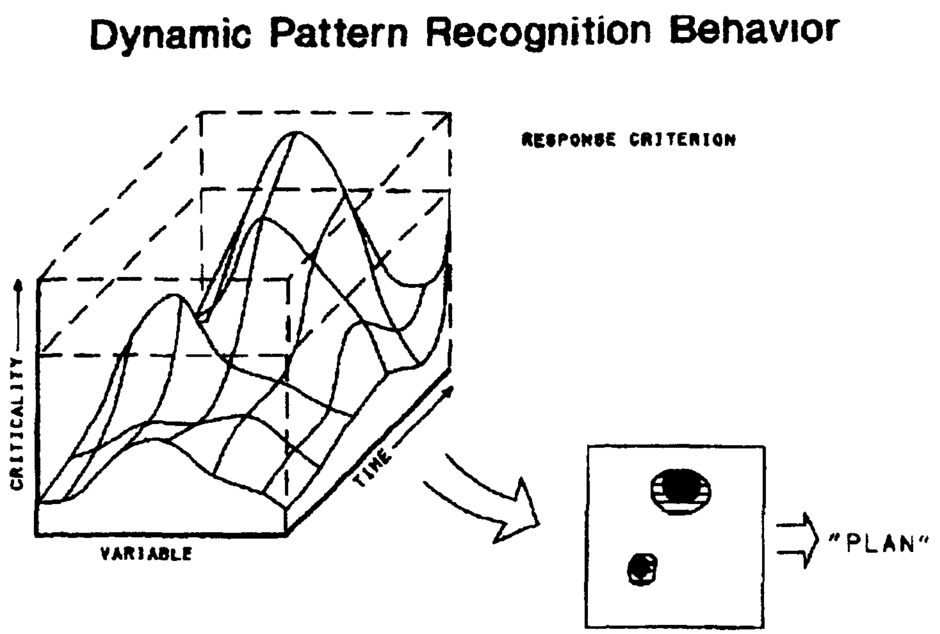


Figure 1. Dynamic Pattern Recognition of Situational Awareness.

The Parallels Between Situational Awareness For Emergency Management And Other Complex Systems

Emergency response personnel require similar situational awareness skills as operators in other complex systems, such as aerospace systems. For example, fire fighters are faced with similar life and death decision making under stress as the fighter pilot. Another parallel, is that incident commanders in high rise, wild fire and other types of fires, are faced with similar decision making to battlefield commanders concerning logistics, resource management, tactics, personnel, etc. Hi-Rise fires offer a prime example of the parallels to the military environment, which lead to the concept of situational awareness.

It involves both individual tactical situational awareness skills, as well as command and control situational awareness skills. Situational awareness is a fundamental human behavior that applies to all complex systems. Hence, the principles and practices developed to aid situational awareness for other applications can be applied to emergency management situations.

What Can Be Done To Improve Situational Awareness In Emergency Management?

In the wake of Desert Storm, many credit the military's outstanding performance to its use of simulation training. Putting pilots and tank commanders in simulations, which create an

event or duplicate one, provided skills directly transferable to the battlefield. Simulation and training technology has been developed for the military to enhance the acquisition of situational awareness and decision making skills. This same type of training should be given to emergency personnel. This technology can produce simulated incidents and training exercises with no risk to the participants. Simulation offers a way for emergency management personnel to practice responses to emergency events, which in the real world are characterized by their infrequency and variability. It allows for a large number of personnel to participate and receive the benefits of the training. It also provides sufficient adaptability to produce unique scenarios of various types of incidents. The resulting simulation and training technology can be designed to aid in training for individual, small group, and large scale emergency applications.

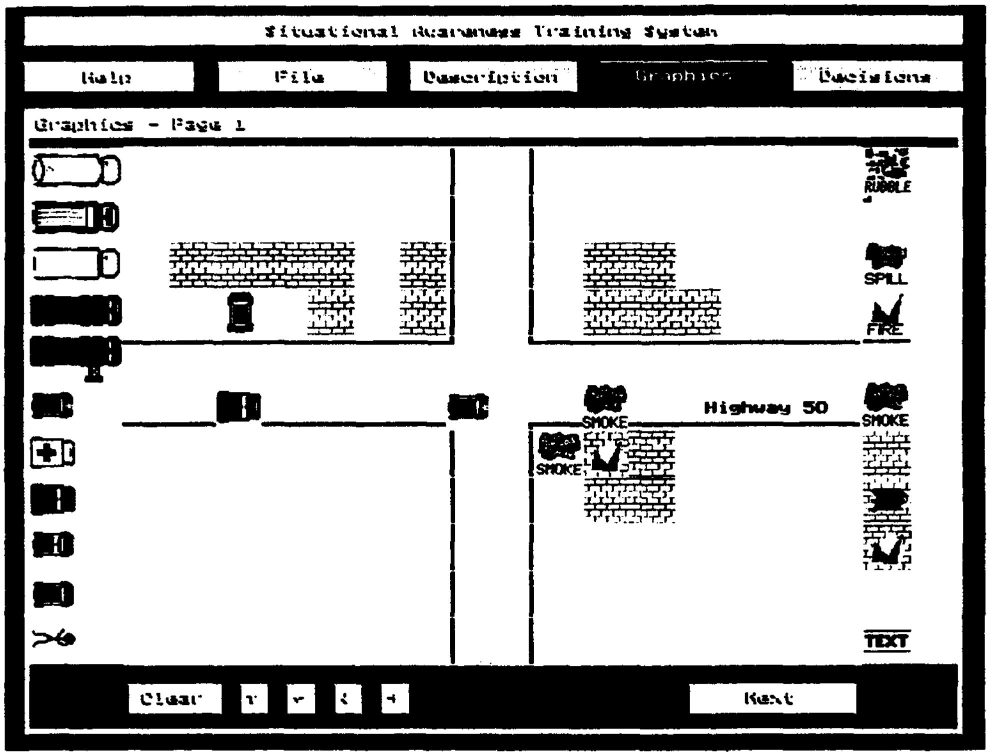


Figure 2. Sample Screen for Situational Awareness Training System.

An Example:

> *Fire is leaping through the roof of a one-story brick building. The district commander orders Engine 33 to stretch a second line into the building. After hearing that the building is empty, the district commander sends in another unit to back up the attack hose lines. Suddenly, the two side walls and rear wall of the building collapse. One fire fighter is dead and three are injured.*

None of the firefighters were hospitalized – this fire took place in a Situational Awareness Training System (Figure 2). This simulation is just a piece of today's high technology solution to emergency-training problems. In the past, emergency personnel have been trained by watching films, reading books, and honing physical skills. When an emergency happened, these trainees encountered situations that required fast decisions under stressful circumstances. An inappropriate decision, while a valuable learning experience, could be disastrous.

Researchers have developed the ability to create simulated emergencies that permit potential nightmares to come alive. The system includes three types of software:

- incident command training system (ICTS)
- situation awareness training system (SATS)
- lessons-learned database

The instructor uses the software to create scenarios or reproduce actual emergencies for simulation. (The programs are designed so that the instructor has to enter only the situational information; the instructor does not have to understand computer programming.) The scenes can be tailored to match actual departmental resources and include buildings that exist within the department's jurisdiction. With this flexibility, a fire department can learn how to handle a fire in the high-rise bank building in their own downtown.

These programs transform a regular PC into a powerful training tool. Skills that are too expensive or too dangerous to train can now be learned inexpensively and without risk. This training method can be effectively applied to other situations, such as emergency medical incidents, treatment of hazardous materials, or airplane crashes. With the proper software and a PC, emergency personnel can be trained simply by booting a computer.

Summary

The primary problems associated with current emergency management procedures, on all levels, originate for inadequate decision making skills, especially in the area of situational awareness. These shortcomings, can easily be generalized to the entire spectrum of emergency management applications. Emergency management, considered as a complex system, has many parallels to other complex systems for which situational awareness has been an intense area of investigation. The understanding of situational awareness skills, both what they are and how they are acquired, emerging from the study of these other systems has direct transfer to emergency management. Given the negative impact of emergency management events on the United States Economy, conservatively estimated between 25 and 30 billion dollars a year, the importance of improving situational awareness skills in this area is clear.

Acknowledgments

This research was partially supported by the State of Florida High Technology and Industry Council.

References

Companion, M. A., Corso, G. M., Kass, S. J., & Herschler, D. A. (1990). *Situational awareness: An analysis and preliminary model of the cognitive process* (IST-TR-89-5). Orlando: University of Central Florida, Institute for Simulation and Training.

Kass, S. J., Herschler, D. A., & Companion, M. A. (1991). Training Situational Awareness Through Pattern recognition in a Battlefield Environment. *Journal of Military Psychology*, 3, 105-112.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. New York: Elsevier.

Situational Awareness In Emergency Services Incident Command

Terry L. Schenk

Seminole County Fire Department

Public safety work in its purest sense is "critical" in nature. Responses to emergency situations include the elements of danger, time crucial action, quick decision making and the compelling need to be right the first time. When the situation involves a large operation such as a high rise building fire, these elements are multiplied: numerous decisions and actions are taking place simultaneously on the scene of the emergency. The most recent example of such a situation was the bombing of the World Trade Center Complex in New York City. Almost all of the designed alarm and assistance systems, as well as the building evacuation plan, failed. The successful outcome of this incident, as measured by the low loss of life, was surely aided by the situational awareness exhibited by emergency personnel, including the people commanding the incident.

Situational awareness plays an important role in critical situations such as this one or any other complex alarm in which an *incident command system (ICS)* must be initiated. While pinpointing specific situational awareness elements in the realm of emergency services is difficult, there are some key characteristics that present themselves. These are: crisis psychology, reactive vs. proactive focus, projected thinking, scope orientation, and alternative option assessment.

When we speak of crisis psychology, we are referring to the "mission importance" attitude that must be developed and emphasized to all participants in the emergency services arena. While much of the preparation work, such as planning and training, is procedural, there is a great deal of critical decision making that must take place in real time at emergency situations.

The first hypothesis I would like to put forward concerns an emergency situation's forced shift in the emphasis from procedures to automatic reaction.

In the aviation field, the pre-flight of an aircraft is very procedural and, in fact, is usually read from a card or manual. An in-flight emergency in many cases (except where the emergency is of a long enough duration to consult a manual) requires an "automatic" response on the part of the flight crew to bring the event to a successful conclusion. In such an *emergency* scenario, situational awareness plays a far more significant role.

Stated simply, "As a situation becomes more "critical" (and less "procedural"), the more dependence there will be on situational awareness." There is an ever present "pressure to act" during emergency responses which, if appropriately followed, can produce a "heroic effort" and if ignored can result in tragedy along with allegations of incompetence.

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

Included in the psychology of crisis response is the ever present thought (by those addressing the situation in a dedicated manner) of the *"impact of the outcome."* In many situations the impact can be matter of life or death! It appears that if an individual is situationally aware in this type of environment, there is a constant (even if only subliminal) reevaluation of the situation and actions being taken. This is not to say that they dwell on their reevaluations or second guess themselves; quite the contrary. To act decisively and to stay ahead of a situation, an incident commander cannot afford to do either of these. The adept, situationally aware incident commander is highly and consistently aware of the "cascading impact" of each decision made.

A second characteristic , therefore, is simply stated as, "The more risk involved in a decision (i.e., the less "safe" it is) the more dependence there will be on situational awareness."

Much public safety work involves *reacting* to stimuli. Dispatchers *react* to calls from the public for service. Emergency response personnel *react* to a buzzer or bell to initiate response to an incident. Upon arrival at the scene, personnel *react* to visual signs, verbal complaints (such as in medical assistance calls) and, in general, the situation as it presents itself. An individual who is situationally aware looks beyond the obvious or what is seen and thinks/acts in a proactive mode. Merely reacting to the obvious manifestations, a situation results in the event being "driven" by what occurs rather than what an incident commander desires to occur. This presents a clear distinction when it comes to situational awareness: those individuals who are identified as being "situationally aware" almost always are those operating in a proactive mode as opposed to a reactive mode.

The third characteristic is that of "projected thinking." Fire/Rescue situations, much like the aviation applications of air combat and air traffic control, are dynamic rather than static. Strategies and decisions based solely on the current situation, rather that what *will be*, are doomed to failure. A situationally aware incident commander must "think ahead" of an incident. This is not just a comfortable way to operate but rather an *imperative* way to operate. The job of the workers or "hands on" personnel at the scene of an emergency is to focus primarily on what needs to be done now. The incident commander needs to focus primarily on what needs to be done 5 minutes, 1 hour, 4 hours, or even, in some situations 2 days from now. It is essential that incident commanders "mentally project" where a situation is going in order to positively influence the outcome. Thus, situational awareness not only involves awareness of *current* conditions but also awareness of *future* possibilities or probabilities.

Recently the most costly disaster in United States history was played out in South Florida. Hurricane Andrew devastated a portion of Dade County, putting all of the emergency services to the extreme test. Critiques of the response to the disaster contained a recurring statement: "We just never imagined that we would be dealing with an emergency of such a large scale." It became apparent that there was quite a misjudgment regarding the potential scope of a major hurricane disaster event. Situationally aware people are oriented to realistically appraising the scope of the situation or possible situation they are, or will be, dealing with.

A fourth characteristic is obtaining the proper scope orientation at the outset of attacking the situation. A common mistake made by the new fire fighter is focusing too much on the visible sign of fire. At a house fire where flames are blowing out of a window on one end of the house, an inexperienced fire fighter will often take a hose stream to that window and flow it into the opening to extinguish the fire. On the surface this seems like an obvious thing to do but, in fact, ignores the *scope* of the problem. The effect of directing the hose stream into the window may be to burn the entire structure down since this tactic pushes the fire towards

the unburned portion of the house. The correct procedure would be to make entrance to the house and advance the hose to the room involved. The stream can then be operated to extinguish the fire and push it out of the window *away* from the unburned portion of the house. A fire fighter who uses this approach is more oriented to the scope of the problem. Situationally aware people are oriented to considering the overall scope of the challenge or problem that faces them.

The fifth characteristic is that of considering and using alternative options. Most of the training that occurs in emergency services tends to focus on making "either/or" decisions. A situationally aware incident commander will look beyond the obvious and derive an alternative option almost from thin air. A company called Aero Innovation has developed a program called WOMBAT (defined as Wondrous Original Method for Basic Airmanship Testing) to test situational awareness and stress tolerance in the aviation environment. Their promotional literature probably provides the best example I have come across to describe the concept of "alternative options." The example is from Okinawa at the end of World War II.

> "A friend taking off in a heavily loaded C-46 lost power from one engine just as the plane became airborne and the end of the runway passed below. Holding altitude caused a rapid loss of airspeed. Advancing the good engine to full power, feathering the prop and closing the cowl flaps on the dead engine and retracting the landing gear were not sufficient. The airman went against the book and closed the cowl flaps on the good engine! This gave him five more knots, just enough to circle and land - with a burned out engine but a safe airplane, crew, passengers and cargo. The airman had defined the aerial problem correctly and seized an unanticipated opportunity."

Situational awareness opens the possibilities of a modified third, fourth or fifth option, and fights against becoming "boxed in" when it comes to decision making.

In observing our incident commanders who appear to be very competent at remaining "situationally aware," I have found three ingredients that seem to be common denominators. These individuals exhibit perceptiveness, confidence and knowledge. In a very base way, these principles were defined by a big bellied sheriff who was once advising a somewhat nervous new city chief of police. The police chief was looking for advice from the tenured sheriff on how to be successful in his new post. The sheriff sat back and said, "Son, if you are going to be successful in a position, you have to do three things: look the part, act the part, know the part, and if you have to, you can get by with any two of the three!" There is no doubt that some of the factors of situational awareness are hidden in this "big bellied sheriffs" axiom.

As researchers continue to study the factors of situation awareness in complex environments, there will no doubt be great improvements in identifying and defining specific characteristics as they relate to unique activities such as emergency response services. When this occurs, it will be interesting to see how much of this will be able to be taught to others to improve individual performance.

Team Awareness

- **The Role of Shared Mental Models in Developing Shared Situational Awareness**
- **Communication and Team Situational Awareness**

The Role of Shared Mental Models in Developing Shared Situational Awareness

Eduardo Salas, Renée J. Stout, & Janis A. Cannon-Bowers

Naval Training Systems Center

Situational awareness is most often brought to the public's attention when it is found to be deficient or lacking. For example, the crew of Eastern Airlines Flight 401 demonstrated a surprising lack of situational awareness in December, 1972, when they became preoccupied with a faulty landing gear indicator light. They failed to recognize that their autopilot had become disengaged, sending them to the ground and to their deaths. In addition, an analysis of 175 military aircraft mishaps attributed to human error revealed that the majority were caused by problems in situational awareness (Hartel, Smith, & Prince, 1991). As a result of such evidence, several authors have indicated the importance of situational awareness for mission effectiveness (cf. Endsley, 1988). While the need for situational awareness is most typically discussed in an aviation context, the importance of situational awareness to mission accomplishment in a variety of operational contexts has been noted as well (Wellens, 1990).

Although its importance to effective performance has been recognized, several conceptual and methodological issues surround research regarding this construct. Included are the following: 1) there is no agreed upon definition of situational awareness, 2) there is a lack of valid and reliable measures of situational awareness, and 3) insufficient attention has been paid to team situational awareness.

This paper addresses the lack of research pertaining to the distinction between individual and team situational awareness. It does so by proposing that cognitive mechanisms (namely, shared mental models among team members) that have been employed to explain team coordination and performance may contribute to the development of team situational awareness. A model is proposed to explain how team situational awareness that may lead to adaptive performance in dynamic, rapidly changing task contexts.

Background

While the importance of teams in the workforce has been well documented (cf. Salas, Dickinson, Converse, & Tannenbaum, 1992), an understanding of what teamwork is and

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

how it relates to team performance has, to a large extent, eluded the research community. Salas et al. (1992) summarized work related to teamwork and team training, and concluded that most efforts have emphasized the training of individual skills in a team setting, rather than training the skills that are demanded by interaction requirements of the team task. Furthermore, Cannon-Bowers, Salas, & Converse (in press), indicated that the majority of the empirical investigations on team processes have essentially ignored the more elusive components of teamwork, such as adaptability and coordination of action, and have instead focused on more readily observable components, such as communication. Furthermore, as noted previously, little attention has been given to how team interactive processes affect team situational awareness (Endsley, 1991a).

Shared Mental Models and Team Performance

A construct that may help to provide an understanding of the relationship between team processes and situational awareness is shared mental models. The concept of mental models has been employed in the investigation of a variety of topics, including situational awareness, and several definitions of mental models have been posited (cf. Rouse & Morris, 1986). While the majority of the literature on mental models has focused on individual cognitive functioning, acquisition of knowledge, and human-system interaction, researchers have also employed the construct of mental models as an explanatory mechanism for coordination in teams. Specifically, in an attempt to address the overlooked aspects of teamwork, recent researchers have employed the construct of shared mental models to explain how team members are able to anticipate and predict each other's needs and thus adapt to task demands and coordinate their activities. It has been hypothesized that shared mental models among team members are required for effective team coordination (Cannon-Bowers & Salas, 1990; Cannon-Bowers et al., in press; Orasanu, 1990; Rouse, Cannon-Bowers, & Salas, 1992). In particular, when rapid, adaptive performance is required in a team, researchers have argued that effective team members draw on common knowledge bases (mental models) that enable them to develop accurate expectations for their own performance and for the performance of teammates. These expectations allow team members to coordinate effectively without the need for extensive overt strategizing.

A question that arises when discussing shared mental models among team members is, "What aspects of the mental model do team members need to share." One approach to addressing this issue is to describe the different types of knowledge that can be shared among team members in performing complex team tasks. For example, it has been proposed that successful teams must share common knowledge of several factors, including: overall task and team goals, individual tasks, team member roles, and the team members themselves (Cannon-Bowers et al., in press).

Taking this notion one step further, Converse and Kahler (1992) conducted an extensive literature review, and proposed that there are three types of knowledge that can be shared among team members: declarative, procedural, and strategic. These types of knowledge incorporate the knowledge types discussed by Cannon-Bowers et al., and serve as a useful categorization system when discussing shared mental model knowledge across a variety of task contexts. According to Converse and Kahler, "declarative models contain information about the concepts and elements in the domain, and about the relationships between them" (p. 5). They contain knowledge of facts, rules, and relationships, and

include knowledge of: the overall system, task goals, the relation among system components, equipment/hardware, positions/roles, and the team members themselves. Procedural mental models "store information about the steps that must be taken to accomplish various activities, and the order in which these steps must be taken" (Converse & Kahler, 1992, p. 6). Procedural knowledge is essentially a sequential and timing type of knowledge which also includes an understanding of task action/goal relationships, as well as how other member actions affect this relationship.

Both declarative and procedural knowledge are consistent with Rasmussen's (1986) behavioral taxonomy (i.e., knowledge-based, rule-based, skill-based), in that they each contain both knowledge-based and rule-based components. The categories of declarative and procedural knowledge go a step further than Rasmussen's taxonomy, however, by identifying that aspects of knowledge must be shared among team members. That is, team members must share portions of their declarative and procedural knowledge bases. Exactly which portions of these knowledge bases must be shared is difficult to determine without empirical investigation; however, some aspects of the model seem more necessary to share than others on intuitive grounds. For example, team members probably need not share detailed information regarding the operation of each other's equipment. However, it would seem that they must understand what information a fellow team member needs from them in order for the team member to complete his/her task, and that they must understand when in a task sequence this information should be presented. An example might help to clarify this point. In a modern fighter aircraft, the pilot flies the plane and the Radar Intercept Officer (RIO) operates the weapons system. While the RIO need not know how to fly the plane, he must know what kind of information the pilot needs to receive in a given engagement (e.g., target heading, speed, altitude, crossing angle).

A final type of mental model proposed by Converse and Kahler (1992) is the strategic model. Strategic mental models "are comprised of information that is the basis of problem solving, such as action plans to meet specific goals, knowledge of the context in which procedures should be implemented, actions to be taken if a proposed solution fails, and how to respond if necessary information is absent" (Converse and Kahler, 1992, p. 6). According to Converse and Kahler (1992), strategic knowledge is a compilation of declarative and procedural knowledge. Through experience, strategic knowledge allows for automatic performance and enables expert team performance via the application of appropriate task strategies.

Strategic knowledge is consistent with Rasmussen's skill-based category. Thus, it is only when team members share strategic knowledge that consistent effective team performance can be accomplished. Using the previous example provided for shared declarative and procedural knowledge, an understanding of team member information requirements and an understanding of when these requirements are crucial in completing a mission, are only generally and globally possessed in declarative and procedural mental models. This knowledge is static, while strategic knowledge is dynamic and is updated based upon mission parameters and team member interactions in response to mission events. Thus, the strategic model takes knowledge stored in declarative and procedural models and applies that knowledge within a dynamic mission complex. As a result, it provides an understanding of: cue/action sequences, cue patterns and their significance, team resources and capabilities, and appropriate task strategies. It can be hypothesized that team members must share information on each of these factors in order to adapt and interact effectively in complex, changing task environments.

General support for the importance of shared mental models among team members has been provided indirectly from the results of several studies (Cream, 1974; Dawes, McTavish, & Shaklee, 1977; Hammond, 1965; Hemphill & Rush, 1952; McIntyre, Morgan, Salas, & Glickman, 1988; Oser, Prince, Morgan, & Simpson, 1991), post hoc to explain results of other studies (Orasanu, 1990; Kleinman & Serfaty, 1989); and via empirical investigation (Adelman, Zirk, Lehner, Moffett, & Hall, 1986; Brehmer, 1972; Volpe, Cannon-Bowers, Salas, & Spector, 1993). Essentially, some support has been found that shared mental models can be used as an explanatory mechanism for implicit coordination (i.e., coordination in the absence of explicit overt strategizing), maintenance of performance under stress or high workload, and the ability of team members to anticipate and predict each other's needs. Recently, Rouse et al., (1992) have proposed that the underlying mechanism of how shared mental models allow all of these things to occur is that they provide the team with shared explanations that lead to shared expectations for performance. In addition to support for the importance of shared mental models among team members to team performance, support for the importance of team situational awareness to team performance has been provided (Brannick, Prince, Prince, & Salas, 1992; Mosier & Chidester, 1991; Orasanu, 1990; Stout, Carson, & Salas, 1992).

Summarizing the results of both literature related to shared mental models and performance and team situational awareness and performance, the following conclusions can be made: 1) mental models are important for individual situational awareness; 2) shared mental models, which allow each member to form adequate explanations and expectations of task and team actions, are important to team performance; and 3) team situational awareness is important to team performance. Expanding upon the above conclusions, it would seem obvious that shared mental models are also important to team situational awareness. Indeed, Endsley (1991a) proposed that the level of situational awareness obtainable for a team can be represented by the overlap of each member's individual level of situational awareness.

Sarter and Woods (1991) suggested, however, that having an adequate mental model of a situation is a necessary prerequisite, rather than a guarantee, that situational awareness will be attained. According to these authors, situational awareness refers to "a continuously updated and integrated mental representation of a mission" (p. 23). Extending this logical argument to a team context, it would seem that shared mental models among team members provide only a necessary prerequisite for achieving shared situational awareness. The dynamics of the task situation would seem to dictate that the shared knowledge bases must be differentially utilized and must be continually updated. In order to begin to understand this process, it is necessary to delineate how shared information among the team members is transformed into shared team situational awareness in a dynamic task situation.

A Model of Shared Situational Awareness

Figure 1 attempts to illustrate the process of how team shared mental models are transformed into team situational awareness. According to the overall flow proposed in Figure 1, it is theorized that effective teams contain members who enter the team setting

with shared pre-existing knowledge bases. That is, effective teams have members whose understanding of the overall mission (declarative knowledge), members, roles, and tasks to meet mission requirements (declarative knowledge), and general sequence of task activities necessary to perform efficiently (procedural knowledge), are compatible. This understanding is necessary, but not sufficient to enable teams to adapt to changing task demands. Most important for effective performance is shared strategic knowledge, which is composed of shared declarative and shared procedural knowledge. In other words, team members must also have a compatible understanding of the context in which they are operating, the actions to be taken when unexpected events occur, and the information that must be obtained in order for them to make critical decisions and take appropriate actions.

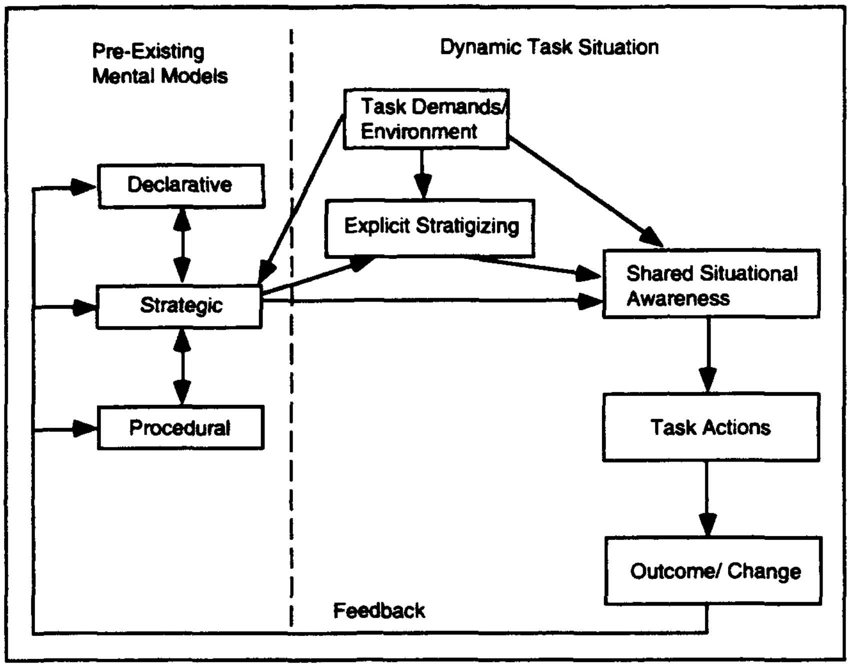


Figure 1. Relationship between shared knowledge and shared situational awareness.

Figure 1 predicts that, given a particular operational context and a given set of mission parameters and events (including external or environmental task factors such as the degree of time pressure that exists and the severity of stressors present), information that is shared in strategic mental models allows team members to: 1) have common explanations of the meaning of task cues, 2) make a compatible assessment of the situation, and 3) form common expectations of additional task and information requirements. This shared level of situational awareness allows them to take appropriate actions, whether gathering additional information critical to making a decision, or implementing a particular procedure. Thus, utilizing pre-existing shared strategic knowledge, team members faced with a dynamic performance situation, must have a common understanding of cue/action sequences, cue patterns and their significance, team resources and capabilities, and

appropriate task strategies. Common interpretation of cues allows for action that is both accurate and expected by teammates.

In task environments where external factors limit the extent to which team members can overtly strategize with each other, the model proposes that shared situational awareness allows team members to engage in "seamless" or implicit coordination. Members are enabled to anticipate and predict how each other member will assess the situation, given a set of cues and events, and are able to take appropriate action in response to task parameters, without the need for explicit strategizing. Members are also provided the opportunity to act in proper sequence via prediction of other's actions as well.

When discussing team settings in which overt strategizing is not severely restricted, or team settings in which such strategizing can occur freely and openly, an interesting dimension is added to the model. Strategizing can help teams develop shared knowledge bases both prior to engaging in their missions and during the missions (i.e., when teams make long term plans prior to their mission and short term contingency plans during their mission; these plans help teams understand what each member will be responsible for, and what each member expects given a set of events in the mission).

In addition, shared knowledge structures can make team strategizing more efficient. First, strategizing can be more efficient when team members have the necessary resources, due to rapid assessment of situational cues via their strategic knowledge, to maintain situational awareness and engage in overt strategizing simultaneously. Results of an effort by Brannick et al. (in press) on team situational awareness support this contention. They found that experienced pilot instructors were able to maintain high levels of performance and to engage in overt verbalizations related to situational awareness in a high workload scenario, while inexperienced students were not. This was presumably because instructors were able to automatize components of performance, enabling a more efficient attainment of situational awareness. Second, strategizing can be made more efficient, since shared knowledge structures permit accurate expectations to be set, and team members can provide needed information in advance of requests for the information, when mission events occur. Results of an effort by Orasanu (1990) on team situational awareness support this portion of the model. She found that effective crews had higher levels of situational awareness and provided necessary information in advance. According to Figure 1, once an action is performed, its impact is evaluated to determine whether the intended outcome was attained. Further assessment of the situation and potential additional necessary actions must be determined. This adaptive performance is enabled when feedback from previous actions updates existing team mental models. Given shared explanations and expectations throughout this cycle, updated information should remain compatible and should lead to compatible assessments of the situation throughout the performance of the mission. Shared strategic knowledge bases that team members bring to the operational context allow for sustained shared team situational awareness, when appropriately updated, based upon mission parameters and team member interactions in response to mission events.

This continual reassessment, given a shared understanding of "the big picture," allows for "fine tuning" of performance. Results of an effort by Mosier and Chidester (1991) on team situational awareness support that effective teams engage in this "fine tuning" behavior. The model proposes that the dynamic reassessment can take place, since updated shared strategic knowledge permits teams to know what additional information to look for to maintain a valid assessment of the changing situation.

Summary

This paper has argued that team effectiveness is a function of (among other things) the team's level of shared situational awareness. Further, it was demonstrated that the construct of shared mental models may help explain how teams can develop and maintain shared situational awareness. Specifically, it was contended that effective teams draw on shared declarative, procedural, and strategic mental models in order to form shared situational awareness. When the environment restricts the team's ability to strategize overtly, this shared situational awareness provides a basis for quick, seamless, implicit coordination. In cases where overt strategizing is less restricted, team members can engage in rapid, efficient strategizing, leading fairly quickly to task action. In still other cases, when strategizing can be extensive, effective teams engage in behaviors that build common knowledge about the task, thereby preparing them to cope with difficult task contingencies. Finally, across different cases, the importance of feedback (regarding the outcome of the action or other team member's behavior) for refining and making team member mental models more accurate, was noted.

References

Adelman, L., Zirk, D. A., Lehner, P. E. Moffett, R. J., & Hall, R. (1986). Distributed tactical decision making: Conceptual framework and empirical results. *IEEE Transactions on Systems, Man, & Cybernetics,* 16, 794-805.

Brannick, M. T., Prince, A., Prince, C., & Salas, E. (1992). *Team coordination: Measuring an elusive construct.* Unpublished manuscript.

Brehmer, B. (1972). Policy conflict as a function of policy similarity and policy complexity. *Scandinavian Journal of Psychology*, 13, 208-221.

Cannon-Bowers, J. A., & Salas, E. (1990, April). *Cognitive psychology and team training: Shared mental models in complex systems*. Symposium presented at the Fifth Annual Conference of the Society for Industrial and Organizational Psychology, Miami, FL.

Cannon-Bowers, J. A., Salas, E., & Converse S. A. (in press). Shared mental models in expert team decision making. In N. J. Castellan, Jr. (Ed.), *Current issues in individual and group decision making*. Hillsdale, NJ: Lawrence Erlbaum.

Converse, S. A., & Kahler, S. E. (1992). *Knowledge acquisition and the measurement of shared mental models*. Unpublished manuscript. Orlando, FL: Naval Training Systems Center.

Cream, B. W. (1974). A functional integrated systems trainer for individual and crew coordination training. *Proceedings of the Fourth Annual Symposium on Psychology in the Air Force.*

Dawes, R. M., McTavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior common dilemma situation. *Journal of Personality and Social Psychology*, 35, 1-11.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society 32nd Annual Meeting*, 97-101.

Endsley, M. R. (1991a). *Situation awareness in dynamic human decision making: Theory.* Unpublished manuscript.

Hammond, K. R. (1965). Essays on research approaches to potential threats to peace. *Journal of Social Issues*, 21, 44-66.

Hartel, C. E., Smith, K., & Prince, C. (1991). *Defining aircrew coordination: Searching for mishaps for meaning*. Paper presented at the Sixth International Symposium on Aviation Psychology, Columbus, OH.

Hemphill, J. K., & Rush, C. H. (1952). *Studies in aircrew composition: Measurement of cross training in B-29 aircrews.* Ohio State University: (AD Number: B958347), Columbus Personnel Research Board.

Kleinman, D. L., & Serfaty, D. (1989). Team performance assessment in distributed decision making. In R. Gilson, J. P. Kincaid, & B. Goldiez (Eds.), *Proceedings of the Interservice Networked Simulation for Training Conference* (pp. 22-27). Orlando, FL.

McIntyre, R. M., Morgan, B. B., Jr., Salas, E., & Glickman, A. S. (1988). Teamwork from team training: New evidence for the development of teamwork skills during operation training. *Proceedings of the 10th Annual Interservice/Industry Training Systems Conference* (pp. 21-27). Washington, DC: National Security Industrial Association.

Mosier, K. L., & Chidester, T. R. (1991). Situation assessment and situation awareness in a team setting. *Proceedings of the 11th Congress of the International Ergonomics Association*, Paris.

Orasanu, J. M. (1990). *Shared mental models and crew decision making* (CSL Tech. Rep. No. 46). Princeton, NJ: Princeton University, Cognitive Sciences Laboratory.

Oser, R., Prince, C., Morgan, B. B., Jr., & Simpson, S. S. (1991). *An analysis of aircrew communication patterns and content* (NTSC Technical Report No. 90-009). Orlando, FL: Naval Training Systems Center.

Rasmussen, J. (1986). *Information processing and human machine interaction: An approach to cognitive engineering*. New York: Elsevier.

Rouse, W. B., Cannon-Bowers, J. A., & Salas, E. (1992). The role of mental models in team performance in complex systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 22, 1296-1308.

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, 100, 349-363.

Salas, E., Dickinson, T. L, Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 3-29). Norwood, NJ: Ablex.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness in the advanced commercial aircraft cockpit. *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 21-25). Santa Monica, CA: The Human Factors Society, Inc.

Stout, R. J., Carson, R., & Salas, E. (1992, August). *Can team process enhance performance?* Paper presented at the Annual Meeting of the American Psychological Association, Washington, DC.

Volpe, C. E., Cannon-Bowers, J. A., Salas, E., & Spector, P. E. (1993). *Walking in each other's shoes: The impact of cross-training on team coordination, communication and task performance.* Unpublished manuscript.

Wellens, A. R. (1990, August). *Group situation awareness and distributed decision-making*. Paper presented at a symposium on "Natualistic Decision-Making" during the 1990 Annual Convention of the American Psychological Association, Boston, MA.

Communication and Team Situational Awareness

Clint A. Bowers, Curt Braun, & Paul B. Kline

University of Central Florida

Introduction

Researchers interested in performance within complex systems have frequently invoked the construct of situational awareness (SA) to represent the operator's knowledge of his current status relative to his desired status. The failure to maintain appropriate SA has been cited as a major causal factor in mishaps involving complex systems (Endsley, 1988). The construct of situational awareness appears especially useful in dynamic situations such as aviation, in which task demands clearly require the operator to monitor his position in three-dimensional space, to react quickly to unanticipated changes in system status, and to alter course as necessary to accomplish the mission. Because the amount of information required to maintain SA is presumed to be greater than can be held in working memory, the maintenance of SA implies a process in which the operator must frequently sample the environmental and system data to infer his position.

The dynamic nature of SA is typically emphasized in the definition of the construct. For example, Sarter and Woods (1991) have defined situational awareness as "the accessibility of a comprehensive and coherent situation representation which is continuously being updated in accordance with the results of recurrent situation assessments (p. 52)." They emphasize that SA must be considered in an appropriate temporal context. That is, that the importance of cues in the environment might vary as a function of time. For example, a relatively routine warning might escalate into a major emergency if it remains undetected for a sufficient period of time.

It also seems important to begin to consider SA in the context of multi-operator systems. As systems grow increasingly complex, more than one operator is required to manage the task requirements. However, recent events (i.e., the Eastern 401 accident) indicate that increasing the number of operators does not always result in high levels of situational awareness. In fact, it might be argued that team situations impose additional demands for situational awareness, such as the demand to be aware of the behaviors and needs of other workers. Although the problem of situational awareness in team performance has received relatively little attention by researchers, the construct of SA might be critical in understanding effective team performance and in designing training interventions to facilitate performance.

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

Situational Awareness in Teams

The extension of SA to multi-operator performance is not novel. In fact, the need to consider other crew members is included, at least implicitly, in even early discussions of the construct. For example, in one of the early discussions of situational awareness, Bolman (1979) provided specific recommendations for maintaining SA in aircraft. These were: crews must be informed of information regarding position, crews must confirm information and resolve conflicts, crews must provide information regarding behavioral alternatives, and crews must share workload. Thus, SA is presented as a crew-level process which relies upon adequate communication for effective performance.

More recently, a recent manual regarding the training of SA has also highlighted the importance of teamwork in maintaining SA within a flight crew (Schwartz, 1990). Schwartz asserts that although each individual in the cockpit has their own level of situational awareness, the group's situational awareness is limited by the working memory of the individuals and must be formed and updated via appropriate communications from each operator. Thus, the quality of a crew's SA is believed to be a direct consequence of the crew's communication ability.

Despite the prevalence of group process issues in the conceptualization of situational awareness, the majority of researchers in SA have focused on the information processing demands imposed upon individuals (i.e., Endsley, 1988; Fracker, 1988, Kass, Herschler, & Companion, 1990; Sarter & Woods, 1991). In fact, there is very little research concerning SA in multi-operator systems. The only hypothesis that has received empirical attention thus far is the position that team SA is related to communication. These studies are summarized in the following section.

Communication and Team SA

An early study of the relationship between communication and team SA was conducted by Orasanu (1990) as part of her research in crew decision making. She analyzed the communications of flight crews during a simulated mission which required crews to diagnose a problem with the aircraft and alter their flight plans accordingly. She found that higher performing crews on this task made a significantly greater number of situational awareness statements than did poor crews. A subsequent study by Orasanu and Fischer (1991) analyzed the relationship between communications and flight performance across two separate aircraft: the Boeing 737 and the Boeing 727. Crews from each aircraft were asked to complete a simulated scenario which included poor weather, system failures and an aborted landing. The results of this study indicated that the frequency of situation awareness statements distinguished between good and poor crews in the B-727, but not in the case of the B-727.

Mosier and Chidester (1991) also attempted to identify a link between communication and situational awareness using an approach similar to that used by Orasanu (1991). They reasoned that the relationship between SA and communication might be most apparent during emergency situations. Therefore, they investigated information transfer during two simulated emergency situations. They found that the number and type of communications was, in fact,

related to crew performance. Further, they noted that effective crews collected information before and after decisions were made, indicating that problems were not "forgotten" after being solved.

The present study attempted to extend our knowledge of team SA by using communications pattern analyses to investigate differences between high- and low-performers on a simulated flight task which was designed to demand high levels of SA. It is hypothesized that crews which demonstrate a consistent pattern of communication will also perform better on the simulated flight task. Furthermore, analysis of the specific sequences within the crew's communication should indicate a more efficient, logical pattern in the effective crews.

Method

Participants

Twenty IFR rated pilots from Embry Riddle Aeronautical University served as subjects for the current study. The qualifications of the subjects included IFR, CFI, CFII, MEI, and ATP. Individual subjects were paid $50.00 for their participation. The 20 pilots formed 10 two person crews.

Apparatus

The low-fidelity flight simulation, Flight Simulator 4.0 was employed for this investigation. The software presented both an instrument panel and external visual scenes on a 13" diagonal, VGA computer monitor. The software operated on a 80386, 33 mhz IBM compatible personal computer. The software was configured to simulate the characteristics of an early model Cessna 210 aircraft. Communications and navigation equipment included one communications radio, one VOR navigation radio, an ADF, clock, and full compliment of attitude, and airspeed instruments. The equipment was configured to allow the study of team processes by following the Guidelines described by Bowers and his colleagues (Bowers, Salas, Prince, & Brannick, 1992).

Simulated Flight Scenarios

Each crew completed two phases of the experiment. The first phase entailed 10 flights, one practice and nine routine flights. Phase 2 entailed 5 additional novel task flights. Routine flights required crews to fly to a specified point under moderate levels of workload (imposed using wind, weather, etc.). The high-SA novel task required crews to perform a simulated reconnaissance mission which required crews to fly at low altitudes and identify buildings on the ground.

The experimenter played the role of all the controllers (e.g., FSS, ground, tower, departure, approach, and center). The experimenter followed a predefined dialogue for each of the flights. This dialogue was followed strictly except when conditions warranted a deviation

from the dialogue (e.g., lost crew in non-radar environment). All Federal Aviation Regulations (FARs), communication and IFR procedures were incorporated into each of the flights.

Communications Analyses

Videotaped recordings of crew interaction were coded using eight coding categories: response, acknowledgment, planning, uncertainty, action, factual, and non-task related. One additional code identifying communications with ATC was also used, but this code was not used in the analyses, as it did not assess coordination between crew members.

The eight types of statements that served as the coding scheme for the study are listed below, along with their respective definitions.

Response Statements

A response was defined as an offering or transmission of information following an uncertainty or action statement. Such statements are characterized by the amount of information provided. Statements that exceeded one bit of information were categorized as responses.

Acknowledgment Statements

An acknowledgment was defined as a statement that affirms the receipt of a command or instruction. Such statements provided no more than one bit of information. This included statements that were binary in nature, such as yes/no, affirmative/negative, and the like.

Planning Statements

A planning statement referred to those statements that communicated the expected or anticipated status of the aircraft or mission. Unlike action statements, statements of intent were not directed toward a particular individual. Rather, such statements were directed toward informing the crew of an upcoming event, plans, goals, etc.

Uncertainty Statements

This classification of statements encompassed all statements of uncertainty. This included: (1) obvious queries for information; (2) indirect questions; and (3) statements intended to elicit information yet were not specifically stated as such.

Action Statements

The action statement was, for the most part, synonymous with a command. Unlike a command, however, the action statement was not confined to the pilot flying. In fact, action statements were also issued by the pilot not flying.

Statements of Fact

A statement of fact was defined as those statements which verbalize obvious realities of the surrounding environment. Such statements included "There is a northwest 'L'", "We are over the NDB outbound", and "We've been one minute outbound on the procedure turn".

Non-Task Related Statements

Non-task related statements were those that had no bearing on the activities of the flight. These included communications such as joking humor, dinner plans, social activities, etc.

Results

Communication rates were computed by dividing the frequency of each communication element by the duration of the flight. Crews were divided into high- and low-performance groups using a median split (the middle two teams were excluded from this analysis). The difference in communication rates between the two performance groups was evaluated using *t*-tests. The results indicated no significant differences between the two groups on any of the communication categories.

In order to assess the role of communications patterns, the communications of the best- and worst-performing teams were analyzed. Markov Chain analyses (Gottman & Roy, 1990) were conducted to assess whether there was homogeneity within the communication patterns of each crew. The results indicated that there were patterns in the communications of each team that occurred more frequently than would have been predicted by chance ($X^2 = 237.25$ and 375.67, respectively, $p < .05$). The data from each team were then combined to test whether the overall communication patterns were homogeneous. The results indicated that there was no homogeneous pattern within the combined data ($X^2 = 82.8$, $p <$ n.s.). Thus, the patterns observed in good and poor teams is not identical, but there did appear to be homogeneous patterns within each team.

Lag-sequential analysis was used in order to assess the prominent communication patterns within the two groups. The results indicated that, in the best team, three distinct patterns occurred at statistically significant levels. First, statements of uncertainty were followed by responses. Second, responses were followed by planning statements. Finally, planning statements were followed by action communications. For the poor team, only two such patterns emerged. The uncertainty-response pattern was also salient in this group. However, in this team responses were followed more frequently by statements of fact.

Discussion

The role of communication in team situational awareness is implied in the very definition of the construct. However, the existing research has not substantially clarified the role of communications in this process. At some level, this might be attributed to the reliance on the analysis of communications frequencies alone. In this study, for example, the analysis of

simple communication frequencies reveals no differences whatsoever between effective an ineffective teams.

The study of communications patterns might represent a useful alternative to researchers in this area. By investigating the flow of information, we might better understand the nature of situational awareness in teams. For example, the present data suggest that one factor that might contribute to the level of team SA is the expression of a plan after receiving a response. This intuitive, clear pattern of communication might serve to quickly orient the team to there present state and provide a clear expectation for future states. It is also interesting to note that actions, which typically represent commands, were the most frequent communications to follow plans. Thus, the effective team appeared to utilize an efficient pattern to assess, plan, act, and then reassess their situation.

The poor team demonstrated a less clear pattern of communications. Like the good team, questions were followed by a response. However, these responses tended to be followed by additional information. No other significant patterns emerged. Thus, in the poor team, the response seemed to "get lost" in the additional data that tended to follow this team's responses. The role of the response in the formation of the team's plan, therefore, is obscured. One might hypothesize that training to improve the efficiency of communication would result in better performance for this crew.

These results indicate that pattern analysis of communications is a promising tool for understanding the nature of SA in teams. However, it should be noted that these results are based on a very small sample size flying one relatively unique type of mission. Future research should assess the degree to which these results generalize across aircraft and mission requirements.

References

Bolman, L. (1979). Aviation accidents and the "Theory of the Situation." In G. E. Cooper, M. D. White, and J. K. Lauber (Eds.), *Resource Management in the Cockpit.* Moffett Field, CA: NASA Ames Research Center.

Bowers, C.A., Salas, E., Prince, C., & Brannick, M. (1992). Games teams play: A methodology for investigating team coordination and performance. *Behavior Research Methods, Instruments and Computers, 24,* 503-506.

Endsley, M. R. (1988). Design and evaluation for situational awareness enhancement. *Proceedings of the 32nd Annual Meeting of the Human Factors Society,* pp. 97-101. Santa Monica, CA: Human Factors Society.

Fracker, M. L. (1988). A theory of situation assessment: Implications for measuring situation awareness. *Proceedings of the 32nd Annual Meeting of the Human Factors Society,* pp. 102-106. Santa Monica, CA: Human Factors Society.

Gottman, J. M. & Roy, A. K. (1990). *Sequential analysis: A guide for behavioral researchers.* New York: Cambridge University Press.

Kass, S. J., Herschler, D. A., & Companion, M. A. (1990). Are they shooting at me: An approach to training situational awareness. *Proceedings of the 34th Annual Meeting of the Human Factors Society,* pp. 1352-1356. Santa Monica, CA: Human Factors Society.

Mosier, K. L. & Chidester, T. R. (1991). Situation assessment and situation awareness in a team setting. *Proceedings of the 11th Congress of the International Ergonomics Association.* Paris: International Ergonomics Association.

Orasanu, J. (1990). *Shared Mental Models and Crew Performance.* Cognitive Science Laboratory Technical Report No. 46. Princeton, NJ: Princeton University.

Orasanu, J. M. & Fischer, U. (1991). Information transfer and shared mental models of decision making. In R. S. Jensen (Ed.), *Proceedings of the Sixth International Symposium on Aviation Psychology,* Columbus, OH: The Ohio State University.

Sarter, N. B. & Woods, D. B. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology, 1,* 45-57.

Schwartz, D. (1990). *Training for Situational Awareness.* Houston: Flight Safety International.

Addendum

- Situational Awareness: Some Reflections and Comments

Addendum – Situation Awareness: Some Reflections and Comments

Mica R. Endsley

Texas Tech University

Following the highly varied and sometimes heated discussions and presentations during the conference, I have been struck with a few thoughts I would like to pass on. As SA is a relatively new field of study (not a new issue however), attendees at this conference were forced to grapple with a large number varied approaches, definitions and thoughts on SA. While the importance of providing SA to operators of various types of systems was largely endorsed at the conference (and indeed has received steadily and rapidly growing attention and support in both government and industrial sectors), those tasked with doing something about it are faced with a fair degree of confusion in sorting through and organizing this disarray. In the interest of assisting researchers and practitioners in this effort, I'd like to propose a few things that I think will help to facilitate progress in this matter.

First, I would like to address the question of terminology. Both the term *Situation awareness* and *Situational awareness* have been used in referring to this construct, with somewhat equal frequency. To a large degree, the difference is not important, other than for promoting ease in communications and conducting literature searches. However, for the sake of consistency, we really should pick one term and stick with it. I agree with Charlie Billings in supporting the use of the term *Situation awareness.* Literally this means "awareness of the situation", which is exactly what we are talking about. By adding an "al" to situation we turn it into an adjective modifying the term awareness, which makes it a "type of awareness relating to situations". The difference is largely semantic, but I think the first is clearer and more accurate.

Secondly, there was much discussion about whether we meant product or process in talking about SA. Almost all of the definitions and work in this area have treated the term SA as meaning the *state* of one's knowledge and I think we should stick with this. There is much to be gained from looking at the processes involved in gaining this state and I highly encourage this line of effort. However for the sake of clarity in the literature, I would like to suggest that we explicitly label this differently in our discussion. We can talk about the process of acquiring SA, maintaining SA or loosing SA. Alternately we can choose to use the term *Situation Assessment* as Charlie Billings has advocated. While I have reservations about this term in that it implies an active search for and deliberation of information, which is sometimes but not always present in obtaining SA, I feel that the term would be fairly adequate for this purpose, as long as we properly describe it and clearly differentiate from the product.

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

An issue more important than terminology, however, is making sure we all *mean* the same thing when we use the term SA. If we allow the term to mean something different to everyone who uses it, it will never be more than a buzzword, have little utility, and seriously stymie progress in the area. There have been many definitions thrown out and there were many more presented at the conference. Most of these are loose operational definitions which vary in the words used, their specificity, and their degree of generalizability, but largely are not contradictory. I think we've mostly been saying the same thing in slightly different ways. I would suggest that perhaps it is more important to clearly state what SA is *not*. I feel very strongly that we cannot allow our use of the term to include decision making and performance. These are clearly different stages which have different factors impacting them and which indicate wholly different approaches in dealing with them. If we choose to throw these other terms in with SA, I believe we will have broadened the term too far to be useful. Furthermore, SA is not all of one's knowledge. It is only that portion pertaining to the state of a dynamic environment. Mental models of the nature of the system, rules, procedures, checklists and the like, while very important and relevant to the decision making process, are fairly static knowledge sources that fall outside of boundaries of our term. Similarly, SA needs to be dealt with as a construct separately from others which act to impact it. Attention, working memory, workload, and stress are all related constructs which impact on SA, but which can also be seen as separate from it. If we subsume any of these constructs within the term SA, we will lose sight of their independent and interactive nature.

SA as a state also needs to be seen as separate from the process of achieving it. As Richard Mogford pointed out, there may be many different and equally valid ways of achieving the same SA. If we tie ourselves to emphasizing a particular way in which SA is achieved, (by communication, displays or direct senses, for instance) we can lose sight of this fact and become highly limited in our understanding of both the processes and the product.

While the lack of consensus in some of the work and presentations on SA can sometimes be frustrating, I have come to see this as a good thing and a natural part of the exploratory process. It seems to me that many of the presenters were simply taking a different perspective on the topic. We are all like the blind men, each feeling a part of the elephant and coming up with different descriptions. This difference in perspectives can be a tremendous strength in arriving a more complete description of the animal.

An issue which has also seemingly caused some confusion is how SA can somehow be a construct which is both a characteristic of an individual and of the system design. As SA is a state, it is a product of both the individuals (through their knowledge and capabilities) and of the environment (through the quality of the system design). Means of improving SA through both these avenues is important to our common goal.

Finally, I think we need to make certain distinctions when talking about SA measurement. There appear to be two fairly different things being talked about under this guise. The first deals with a quantification of SA – how much does a person have, how complete, how accurate? And in so measuring, as compared to what? Dick Pew pointed out that we can talk about SA as compared to the ideal (total reality) or as compared to an obtainable ideal (that subset displayed to the operator). Each conveys a very different assessment. I would suggest that if we want to talk about evaluating *systems*, we must compare a person's SA to the ideal or we will constantly run into a ceiling effect in our measurements. Much of our technological achievements will directly alter the obtainable ideal in some way. The addition of better sensors, AI and new display technologies make the obtainable ideal a moving target. If we want to talk about the evaluation of *operators* with a stable system design, however, using the obtainable ideal as the benchmark comparison point may be entirely appropriate.

A second very different issue discussed under the label of measurement is a qualitative as opposed to quantitative assessment. That is, these efforts seek to describe SA as a process or in terms of its qualitative nature, as opposed to identifying some exact quantity of it that one possesses. Although not wholly without overlap, I would suggest that these are very different goals to which very different approaches are applicable. I am beginning to see that perhaps some of the differences in opinion on SA measurement expressed at this conference are indeed differences in perspectives between quantitative and qualitative approaches. Again, for the sake of clarity, perhaps we should reserve the term *SA Measurement* for methods of quantifying SA and the term *SA Analysis* or methods of qualitatively describing it.

Finally, I would like to commend CAHFA, specifically Dick Gilson, Dan Garland and Eric Gruber, for managing to put on a high quality conference. They surpassed all expectations in their ability to assemble the diversity and quality of information represented. The tremendous amount of interaction at the conference, both formal and informal, speaks to this accomplishment. Based on my perceptions at the conference and my understanding of the level of cross-fertilization of ideas and information exchanged, I project that we will see a tremendous increase in researcher SA and productivity as a result of their efforts.

Commentary

- **Situational Awareness in Complex Systems: A Commentary**

Situation Awareness in Complex Systems: A Commentary

Charles E. Billings

The Ohio State University

Introduction

This has been a most stimulating and provocative conference. I have learned that situation awareness (SA) is a process – or is it a product? Or is it both, or neither? It either is, or is not, critical to define SA precisely. It either can, or cannot, be quantified, but if it can be, the Heisenberg principle probably applies and we alter it in the process of measuring it. Our sponsors have obviously chosen a topic of some considerable complexity. As Rex Harrison said, "It's a puzzlement!" The structure of the conference implies a search for a general or common framework within which SA can be explored without, as Dick Pew said, being modeled at such a general level that the theory becomes useless as a tool for prediction. I come at this from a somewhat different perspective from most of the participants, having been trained in medicine. It is true that our theorizing has been as seriously handicapped by our relative ignorance of molecular biology, as has psychology's attempts to build adequate theories which have been handicapped by a lack of understanding of central nervous system function. But by and large, the psychologists have done better than we physicians have, by being extremely rigorous in their formulations and by truly elegant experimental approaches.

The phenomenon of situation awareness (I'll beg for the moment the question of whether it is process, product or both) must be a good laboratory psychologist's worst nightmare. As an epidemiologist of sorts, let me suggest why. In its most fundamental terms, SA represents yet one more effort on the part of the human organism to remain distinct from its environment. In Selye's terms, it represents one part of a coping response to environmental stress. Peter Hancock put it very well: "The organism proposes, the environment disposes."

SA is influenced more or less directly by all of the potential stresses in the physical and operational environment. It is influenced very directly by the vehicle or other complex system within which the human is trying to accomplish useful work. As if this weren't enough variance, SA is also influenced by the state, background and knowledge base of the human organism itself, only a portion of whose physical and mental capacity is available for external work (the rest being given over to maintenance of internal homeostasis). It is hardly surprising that we have found it difficult to understand this construct, let alone measure or predict it.

Situational Awareness in Complex Systems
Edited by R. D. Gilson, D. J. Garland, & J. M Koonce

Discussion

There are several "common threads" in what I have heard, and in the literature several have cited in some wonderfully provocative presentations. I return to my epidemiological analogy for help in voicing them. The first is well-honed perceptual senses and capability. Bryce Hartman (1991) says SA is a cognitive phenomenon, and indeed it is—but it is triggered by an "almost magical" perceptual sensitivity in the best fighter pilots, as we have been reminded—and near-threshold perception can be selected for and improved by training. I worry that advanced automation may permit these skills to atrophy in civil aviation, yet there will still be unannounced threats out there, as there always have been!

Given excellent perception and well-developed attentional skills (and these also can be improved by training), we have made the task of deriving a coherent construct of our situation both easier and more difficult. It is easier because integrated displays tell us much that we formerly had to figure out on our own. It is more difficult, because there is much more data from which information must be extracted. Like our colleagues in the nuclear power industry, we have not yet done a very good job of making the right information most salient. I remember an orthopedic surgeon for whom I worked exclaiming during surgery to his long-suffering scrub nurse, "Don't give me what I asked for—give me what I want!" The formation of a coherent construct may not require a great deal of information, but it does require the *right* information. Don't give them what they ask for—give them what they *need!*

Dr. Endsley's elegant summary, and others, made it crystal clear that SA must exist at several levels. I've mentioned her level 1, perception and attention, and level 2, comprehension and integration. Level 3, the ability to predict the near-term future actions of the system controlled, is crucial. Being "ahead of the airplane," not "losing the bubble," are what SA is all about in highly dynamic systems, yet we have done less than we should to provide pilots and controller with trend information that could help them to recognize developing problems before they reach serious proportions. Wiener and Curry (1980) pointed out that such information can assist pilots to develop trust in automated systems. It can also support planning behavior.

The U. S. Air Force Chief of Staff's question about situation awareness, as well as his comment that "I recognize it when I see it," has obviously motivated a major research effort within that Service. The program as described can certainly provide tools and data of great value, yet I am concerned that the holistic design of SAINT, in which combat performance seems to be *the* outcome variable, may preclude a comparison of its results with the results of other SA research.

I find Dr. Endsley's model helpful: situation awareness leads to decision making, which in turn motivates actions to improve the situation. The Air Force studies would be more useful to the community as a whole if they were designed to permit partitioning of those results into these three bins. Decision making and execution are not parts of situation awareness; SA is instead the necessary precursor to effective decisions, which are required for appropriate actions. I would urge that the USAF consider the formation of a steering committee to help insure that the results of its ambitious program takes account of, and proceeds in concert with, the other excellent work on SA ongoing here and in the United Kingdom. I think both the Air Force and the rest of us would benefit.

Let me return to where we began. Lloyd Hitchcock's construct of a "cognitive centroid" is a much more elegant way of saying what I was trying to convey earlier: that each of us puts his or her own "spin" on what is available to be perceived, and to that extent, each of us

develops a unique view of the world around us. The world is the same, but our internal models of it are not. Pilots and air traffic controllers do a lot of reality testing, however – at least the good ones do – and their world models may actually approximate the real world on most occasions.

We must not forget, however, the interesting comments about the quite different world model of the Polynesian navigators offered by Dr. Richard Mogford. It is quite possible that our mental models are unique, and that the decisions taken are likewise unique to some degree. If the execution of those decisions leads to an appropriate outcome, it cannot be faulted simply because it was reached by following a different drummer. Any viable theoretical or experimental approach to situation awareness must recognize that the process may be idiosyncratic, and that my good SA may differ in some respects from yours, just as our backgrounds, knowledge bases, cognitive and decision-making styles may differ.

The problem here is *not*, in my view, our construct of situation awareness. I believe I've heard fairly substantial agreement on what it is and how it happens, though some of us have used different words than others. The cognitive dissonance at this conference has instead been in how to measure SA and in what various measures are trying to tell us about the phenomenon.

The problems associated with modeling the process of situation assessment, which leads to the product, situation awareness, are truly formidable, given the limitations of short-term memory and our limited understanding of the constraints on the chunking process and the cueing processes that allow us to incorporate long-term conscious memory into awareness. These problems are difficult enough in static situations and become far more intractable in a highly dynamic environment. Many of you are far more knowledgeable than I in this domain, and I stand in awe of the elegance of some of the constructs you have presented here.

As an applied scientist, I am particularly concerned about certain aspects of this problem. As you know, I have posited that pilot *involvement* is critical to the pilot's ability to remain in command – aware of and "on top of" the situation (Fig. 1). David Hopkin spoke eloquently about the extent to which ATC automation may lessen the controller's direct involvement in air traffic management. The electronic flight strip issue has the potential to be a really nasty problem, though we can't prove that yet, and by the time we can, the issue may well be moot. We are progressing in a similar direction with aircraft considerations as well, and the rarity of aircraft accidents makes it difficult, again, to *prove* that it is a bad approach. Your situation awareness models suggest to me that we remove active feedback modalities from the human-machine loop at our peril.

PREMISE:
The pilot bears the ultimate responsibility for the safety of flight operations.

AXIOM:
The human operator must be in <u>command</u>.

CORROLARIES:
To command effectively, the human operator must be <u>involved</u>.
To be involved, the human operator must be <u>informed</u>.

Figure 1. First principles of human-centered aircraft automation (adapted from Billings, 1991).

Summary

Let me try to summarize and perhaps simplify some of what I will carry away from this conference. Much of it you have given me, for which I am truly grateful to our speakers and discussants. Some of it represents my own biases, which I shall take delight henceforth in referring to as my own cognitive centroid!

I propose that situation (not situation*al*) awareness is the *product* of situation assessment, a cognitive *process*. The SA gestalt which is formed, and recursively tested, is the basis for primed decision-making. Decisions, whose appropriateness depends on knowledge and rules, are executed more or less effectively depending on the psychomotor skills of the operator. If these are sequential rather than parallel processes, and it seems to me that they must be even though they may be carried out very quickly, then in some cases we may be able to infer decision from action, and gestalt from decision, as long as we understand that the cognitive centroid of the individual operator is idiosyncratic and unknowable to some degree. Training can help to improve perception; it can also help to standardize the decisions taken in a given situation. The comprehension and integration of sensed data can also be improved by training, practice and criticism. The changes brought about by carefully targeted training can be observed and can also help us to understand the underlying processes.

I am not attempting to suggest that the process of situation assessment is so obscure as to be a purely philosophical matter. We know it is a process, we know that it can be learned, and we can either observe (in a more or less intrusive manner) or infer (from decisions) its product. I do suggest that attempts to probe situation awareness directly and immediately are invariably intrusive and that they may often change the level or focus of situation awareness, or even produce responses that bear little resemblance to the actual product or process. Carefully-designed simulations can be of great help in learning more about this product.

What remains for those of us who must apply this knowledge is to insure that the input data are available and are not easily lost in clutter. Given today's information management technology, we can easily drown the operator in data, and we have done so in many advanced applications involving process control. Our most important task, I believe, is to do a more effective job of managing the input data stream in order to simplify the perception and situation assessment process. We must do this in a way that does not cause the operator to lose sight of the system model, lest he or she be unable to recognize that the automated system is failing. This, I think, is the key to enhancing the product, situation awareness, and to improving the decisions taken in difficult situations.

In complex and largely transparent systems, we must always remember that our internal system model is a drastic simplification of the real system, and that this can present the operator with novel challenges when automation begins to fail. There is still a place for simplicity, a verity we seem to have forgotten in our push toward more complex and versatile automation. If the "real" system is simple enough for its dynamics to be grasped by a skilled operator, then the operator's correct internal model of the system will both improve situation awareness and lead him or her much more quickly and certainly to a correct diagnosis if the machine system begins to falter. And that, of course, is one of the major reasons the human operator is there in the first place.

References

Billings, C.E. (1991): *Human-Centered Aircraft Automation: A Concept and Guidelines.* Moffett Field, CA: NASA Technical Memorandum 103885.

Hartman, B.O. & Secrist, G.E. (1991): *Situational awareness is more than exceptional vision.* Aviation, Space, and Environmental Medicine, 62: 1084-89.

Wiener, E.L. & Curry, R.E. (1980): *Flight-deck automation: promises and problems.* Moffett Field, CA: NASA Technical Memorandum 81206.

Approved for public release;
distribution unlimited.

AIR FORCE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12
Distribution in unlimited.
Joan Boggs
STINFO Program Manager